PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/53, 15/82, 9/02, A01H 5/00

A2

(11) International Publication Number: **WO 99/53073**

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/US99/08400

(22) International Filing Date: 16 April 1999 (16.04.99)

(30) Priority Data:

60/081,936	16 April 1998 (16.04.98)	US
60/123,168	5 March 1999 (05.03.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application

US	60/123,168 (CON)
Filed on	5 March 1999 (05.03.99)

(71)(72) Applicants and Inventors: BROUN, Pierre [US/US]; 1249 Capuchino, Burlingame, CA 94010 (US). SHANKLIN, John [US/US]; 4 Duchess Street, Shoreham, NY 11786 (US). WHITTLE, Edward, J. [US/US]; 164 6th Street, Greenport, NY 11944 (US). SOMERVILLE, Chris [US/US]; 5 Valley Oak, Portola Valley, CA 94028 (US).

(74) Agents: KOKULIS, Paul, N. et al.; Pillsbury Madison & Sutro LLP, 1100 New York Avenue, N.W., Washington, DC 20005 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: INTERCONVERSION OF PLANT FATTY ACID DESATURASES AND HYDROXYLASES

(57) Abstract

A method is provided for modifying a fatty acyl desaturase to a fatty acyl hydroxylase consisting of identifying and changing as few as four amino acid residues that are conserved in functionally equivalent desaturase enzymes from various plant species but that are not identical in fatty acyl hydroxylases that exhibit significant overall sequence similarity to the fatty acyl desaturases, and which catalyze hydroxylation at one of the carbon residues on the fatty acyl substrate that is desaturated by the corresponding desaturase; the modifications being made by changing the amino acid residue so that it is identical or functionally equivalent to the amino acid residue found in the naturally occurring hydroxylase. Also provided is a similar method of modifying a fatty acyl hydroxylase to a fatty acyl desaturase by changing seven or fewer amino acid residues. Transgenic plants and products of such transgenic plants wherein the plants have been modified to produce a modified hydroxylase or desaturase are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERCONVERSION OF PLANT FATTY ACID DESATURASES AND HYDROXYLASES

The invention described herein was made in the course of work under grant number DE-FG02-94ER20133 and grant number DE-FG02-97ER20133 from the U.S. Department of Energy. The U.S. government may retain certain rights in this invention.

FIELD OF THE INVENTION

The present invention concerns the modification of nucleic acid sequences and constructs, and methods related thereto, and the use of these sequences and constructs to produce modified enzymes which exhibit altered catalytic activities. The modified nucleic acids are of utility in producing genetically modified plants for the purpose of altering the fatty acid composition of plant oils, waxes, and other fatty acid-containing compounds. Particularly, the present invention concerns the modification of nucleic acids for the selective production of plant fatty acid desaturases and hydroxylases.

BACKGROUND OF THE INVENTION

In addition to common plant fatty acids such as linoleic or linolenic acids, a number of plant species accumulate hydroxylated fatty acids. For example, castor (*Ricinus communis*) accumulates a seed oil which may contain more than 80% ricinoleic acid (Δ-12-hydroxyoctadec-cis-9-enoic acid). This industrial fatty acid, used in the fabrication of lubricants and certain types of nylon, is also present in a number of other unrelated plant species. Biosynthesis of ricinoleic acid in castor has been studied in detail (van de Loo et al., 1995). By a single enzymatic step involving a membrane bound iron-containing enzyme, oleic acid is esterified to phosphatidylcholine. The reaction requires molecular oxygen, NAD(P)H and cytochrome b5 as an electron donor. In all these aspects, oleate hydroxylation has extensive similarity to microsomal oleate desaturation, a key step in the biosynthesis of linoleic acid. Genes encoding the castor and *Lesquerella fendleri* oleate hydroxylases have been identified and the gene products have been shown to have a high degree of sequence similarity to plant oleate desaturases (van de

Loo et al., 1995; Broun et al., 1997).

The castor oleate hydroxylase is about 70% identical to some oleate desaturases, and contains clusters of histidine residues diagnostic of class III diiron-oxo proteins, which include: plant desaturases FAD2, FAD3, FAD5, FAD6, FAD7, FAD8, bacterial alkane and xylene hydroxylases, carotene hydroxylase, carotene ketolase, and sterol methyl oxidases among others (Shanklin et al., 1997). The oleate hydroxylase from the crucifer *Lesquerella fendleri* shows about 81% sequence identity to the oleate desaturase from the crucifer *Arabidopsis thaliana* and about 71% sequence identity to the oleate hydroxylase from the more distantly related species, *Ricinus communis* (Broun et al., 1998). The observation that two different crucifer enzymes are more closely related than the two hydroxylases, and the presence of ricinoleic acid in a small number of distantly related plant species, suggests that the capacity to synthesize ricinoleate has arisen independently several times during the evolution of higher plants, by the genetic conversion of desaturases to hydroxylases.

All higher plants contain one or more oleate desaturases that catalyze the O_2-dependent insertion of a double bond between carbons 12 and 13 of lipid-linked oleic acid ($18:1^{\Delta 9}$) to produce linoleic acid ($18:2^{\Delta 9,12}$) (Shanklin et al., 1998). By contrast, only fourteen species in ten plant families have been found to accumulate the structurally related hydroxy fatty acid, ricinoleic acid (D-12-hydroxyoctadec-*cis*-9-enoic acid) (van de Loo et al., 1993). Ricinoleic acid is synthesized by hydroxylation of oleic acid by enzymes that have similar enzymatic properties and exhibit a high degree of sequence similarity to oleate desaturases (Moreau et al., 1981 and van de Loo et al., 1995). The oleate desaturases and hydroxylases are integral membrane proteins, which are members of a large family of functionally diverse enzymes that includes alkane hydroxylase/alkene epoxidase, xylene monooxygenase, carotene ketolase, and sterol methyl oxidase (Shanklin et al., 1998). Biochemical evidence suggests that these nonheme iron-containing enzymes use a diiron-cluster for catalysis (Shanklin et al., 1997). They contain three equivalent histidine clusters that have been implicated in iron binding and shown to be essential for catalysis for several desaturases (Shanklin et al., 1998). This class of integral membrane proteins exhibit no significant sequence identity to the soluble diiron-containing enzymes which represent a similar diversity of enzymatic activities that include plant acyl-ACP desaturases, methane monooxygenase, propene monooxygenase and the R2 component of ribonucleotide reductase (Shanklin et al., 1998). From the results disclosed herein which demonstrate that amino acid substitution at certain conserved residues in the hydroxylase and desaturase enzymes confer enzymatic function, it is likely that plant oleate hydroxylase genes and desaturase genes are

evolutionarily related.

Most of the plant species that are grown for production of oils do not produce significant amounts of hydroxylated fatty acids. Thus, there is interest in being able to modify oil-producing species so that they produce hydroxylated fatty acids. This may be accomplished by the introduction of genes encoding fatty acyl hydroxylases. Plant genes for fatty acyl hydroxylases from *R. communis* and *L. fendleri* have been described and have been shown to be useful for modifying plants to produce hydroxylated fatty acids (Broun and Somerville, 1997). In addition, methods for using these genes to isolate hydroxylase genes from other plants have been described in U.S. Appln. Nos. 08/530,862 and 08/597,313; and international Appln. Nos. PCT/US95/11855 (WO 96/10075) and PCT/US97/02187 (WO 97/30582), the complete disclosure of which is fully incorporated herein by reference.

In the aforementioned patent applications, we disclosed that an alternative method for the production of hydroxylated fatty acids is to modify a fatty acyl desaturase so that it catalyzes fatty acyl hydroxylation instead of, or in addition to, fatty acyl desaturation. Conversely, since it is also useful to control the degree of fatty acyl unsaturation in transgenic plants by the expression of introduced genes, it is also potentially useful to modify a fatty acyl hydroxylase so that it catalyzes fatty acyl desaturation. Such a modified gene could be used to increase the amount of desaturase activity in a plant.

In order to identify which amino acid residues are responsible for the different catalytic activities of the oleate hydroxylases and the oleate desaturases, the castor and *L. fendleri* oleate hydroxylase sequences are compared herein with the sequences of various oleate desaturases. The concept underlying this comparison was that if a particular residue was conserved in all known oleate desaturases but differed from the castor and *L. fendleri* oleate hydroxylases, it could be important in determining the outcome of the reaction. By contrast, if a particular residue was not conserved among the desaturases, it was unlikely to be responsible for the outcome of the reaction. The results of this comparison indicate that there are only seven amino acid residues which are conserved among all the desaturases but which differ in the oleate hydroxylases. These seven amino acid residues were disclosed in the aforementioned patent applications. Four of the seven critical residues are very close to putative iron ligands suggesting a role for these amino acids in protein function.

Once the amino acid residues of interest have been defined, there are many methods for producing genes encoding modified enzymes, including mutagenesis of existing genes and synthesis of novel genes. The most specific way of obtaining modified enzymes is by

site-directed mutagenesis, enabling specific substitution of one or more amino acids by any other desired amino acid. Site-directed mutagenesis can be performed, after cloning the encoding gene, by mutagenesis in vitro or in vivo and expression of the encoded enzyme by causing transcription and translation of the mutated gene in a suitable host cell.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides novel modified hydroxylase and desaturase enzymes, obtained by expression of genes encoding said enzymes having amino acid sequences which differ in at least one amino acid from the corresponding wild-type enzymes. These mutant enzymes exhibit novel catalytic properties for modifying plant oil composition. A preferred embodiment of the invention is a mutant of the *Arabidopsis thaliana* FAD2 desaturase.

It is a one object of this invention to provide specific amino acid substitutions that result in the conversion of an enzyme which exhibits primarily oleate hydroxylase activity to an enzyme which exhibits more oleate desaturase activity than oleate hydroxylase activity. It is a further object of this invention to provide seven or fewer specific amino acid substitutions that result in the conversion of an enzyme which exhibits primarily oleate hydroxylase activity to an enzyme which exhibits more oleate desaturase activity than oleate hydroxylase acitivity.

It is also an object of this invention to provide specific amino acid substitutions that result in the conversion of an enzyme which exhibits primarily oleate desaturase activity to an enzyme which exhibits more oleate hydroxylase activity than oleate desaturase activity. It is a further object of this invention to provide as few as four specific amino acid substitutions that result in the conversion of an enzyme which exhibits primarily oleate desaturase activity to an enzyme which exhibits more oleate hydroxylase activity than oleate desaturase activity.

It is a further object of this invention to describe a general method by which any fatty acyl desaturase can be converted to a fatty acyl hydroxylase. It is a yet further object of this invention to disclose a general method by which any fatty acyl desaturase can be converted to a fatty acyl hydroxylase.

In another aspect, the invention provides a transgenic plant, comprising a plant that has been modified by the introduction of a gene for a modified hydroxylase.

It is a further object of the invention to provide a transgenic plant, comprising a plant that has been modified by the introduction of a gene for a modified desaturase.

In a further aspect this invention provides a system, which identifies and produces mutant

fatty acyl desaturase, hydroxylase, desaturase/hydroxylase enzymes with novel properties that can be used to modify plant oil composition.

These and other aspects of the invention will be further outlined in the detailed description hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a multiple sequence alignment of deduced amino acid sequences for oleate hydroxylases and microsomal Δ12 desaturases. Abbreviations are: Rcfah12, oleate 12-hydroxylase gene from *R. communis* (van de Loo et al., 1995); LFAH12, oleate 12-hydroxylase gene from *L. fendleri*; Atfad2, FAD2 desaturase from *Arabidopsis thaliana* (Okuley et al., 1994); Gmfad2-1, FAD2 desaturase from *Glycine max* (GenBank accession number L43920); Gmfad2-2, FAD2 desaturase from *Glycine max* (Genbank accession number L43921); Zmfad2, FAD2 desaturase from *Zea mays* (PCT/US93/09987).

Figure 2 shows a general strategy for site-directed mutagenesis of *A. thaliana* FAD2 oleate desaturase.

Figure 3 shows the nucleic acid sequence of the coding region of the *A. thaliana* FAD2 oleate desaturase gene and the corresponding amino acid sequence of the enzyme.

Figure 4 shows a general strategy for introducing seven mutations into the *A. thaliana* FAD2 gene.

Figure 5 shows a comparison of the nucleic acid sequences of the coding regions of the *A. thaliana* FAD2 gene and the mFAD2 gene.

Figure 6 shows a comparison of the deduced amino acid sequences of the *A. thaliana* FAD2 gene and mFAD2 gene.

Figure 7 shows a comparison of the nucleic acid sequences of the coding regions of the *L. fendleri* FAH12 gene and the mFAH12 gene.

Figure 8 shows a comparison of the deduced amino acid sequences of the *L. fendleri* FAH12 gene and the mFAH12 gene.

Figure 9 shows the fatty acid composition of yeast cells expressing desaturase and hydroxylase genes.

Figure 10 shows the genetic complementation of the Arabidopsis *fad2* mutation with the m_7LFAH12 gene.

Figure 11 shows the fatty acid content of seed lipids from independent transgenic Arabidopsis lines expressing m_7FAD2 or m_4FAD2 under control of the *B. napus* napin promoter.

Figure 12 shows the contribution of individual amino-acid substitutions to the activity of the modified *Lesquerella* hydroxylase.

DETAILED DESCRIPTION OF THE INVENTION

One subject of this invention is a class of enzymes, designated fatty acyl hydroxylases, that introduce a hydroxyl group into fatty acids. For example, the fatty acyl hydroxylases of the invention can catalyze hydroxylation of oleic acid to 12-hydroxy oleic acid (i.e., ricinoleic acid) and icosenoic acid to 14-hydroxy icosenoic acid (i.e., lesquerolic acid). This enzyme is referred to herein as "oleate hydroxylase". These enzymes have also been referred to as a class of kappa hydroxylases to indicate that the enzyme introduces the hydroxyl group three carbons distal (i.e., away from the carboxyl carbon of the acyl chain) from a double bond located near the center of the acyl chain.

A second subject of this invention is a class of enzymes, designated fatty acyl desaturases, that introduce double bonds into fatty acids. For example, fatty acyl desaturases of the invention can introduce a double bond between carbons 12 and 13 (counting from the carboxyl end) of eighteen carbon fatty acids. This enzyme is referred to herein as "oleate desaturase".

The above enzymes are named oleate hydroxylase and oleate desaturase in accordance with their discovery in oleate-containing plants. However, we have previously shown that the native enzymes are able to metabolize fatty acids with chain lengths other than eighteen carbons. Similarly, the present invention is not limited to metabolizing oleic acid but can also produce

saturated and/or hydroxylated fatty acids of varying chain lengths. Preferred are substrates with chain lengths of 16, 18, 20 and 22 carbons.

For example, the following fatty acids are also the subject of this invention: palmitoleic acid, hexadec-*cis*-9-enoic ($16:1^{cis\Delta 9}$); hydroxypalmitoleic acid, 12-hydroxy-hexadec-*cis*-9-enoic ($12OH\text{-}16:1^{cis\Delta 9}$); oleic acid, octadec-*cis*-9-enoic acid ($18:1^{cis\Delta 9}$); ricinoleic acid, 12-hydroxyoctadec-*cis*-9-enoic acid ($12OH\text{-}18:1^{cis\Delta 9}$); octadec-*cis*-9,15-dienoic acid ($18:2^{cis\Delta 9,15}$); densipolic acid, 12-hydroxyoctadec-*cis*-9,15-dienoic acid ($12OH\text{-}18:2^{cis\Delta 9,15}$); icosenoic acid ($20:1^{cis\Delta 11}$); lesquerolic acid, 14-hydroxy-*cis*-11-icosenoic acid ($14OH\text{-}20:1^{cis\Delta 11}$); *cis*-11,17-icosadienoic acid ($14OH\text{-}20:2^{cis\Delta 11,17}$); auricolic acid, 14-hydroxy-*cis*-11,17-icosadienoic acid ($14OH\text{-}20:2^{cis\Delta 11,17}$); erucic acid, docos-*cis*-13-enoic acid ($22:1^{cis\Delta 13}$); and hydroxyerucic acid, 16-hydroxydocos-*cis*-13-enoic acid ($16OH\text{-}22:1^{cis\Delta 13}$). It should be noted that icosenoic acid is sometimes spelled eicosenoic acid.

A further subject of this invention is the creation of modified genes encoding modified enzymes which have novel catalytic activities. The term "modified enzyme" as used herein always refers to the product of a modified gene rather than to any direct modification of the corresponding wild-type (WT) protein. Thus, we describe modified enzymes which before genetic modification had desaturase activity but did not exhibit detectable hydroxylase activity, but which after modification exhibit hydroxylase activity. We refer to these enzymes as "synthetic oleate hydroxylases". This designation does not preclude the possibility that the modified enzyme may also retain some amount of desaturase activity.

Similarly, we describe modified enzymes which before modification had oleate hydroxylase activity but had low levels of oleate desaturase activity relative to the amount of hydroxylase activity, but which after modification exhibit higher levels of oleate desaturase activity. We refer to these enzymes as "synthetic oleate desaturases". This designation does not preclude the possibility that the modified enzyme may also retain some amount of hydroxylase activity.

This invention is based on the discovery that plant oleate hydroxylases and oleate desaturases are structurally related enzymes (van de Loo et al., 1995). Indeed, because these enzymes are highly similar in primary structure, we have previously described methods for distinguishing between the two types of enzymes based on a comparison of the amino acid sequences (U.S. Appln. Nos. 08/530,862 and 08/597,313; international Appln. Nos. PCT/US95/11855 (WO 96/10075) and PCT/US97/02187 (WO 97/30582), referred to herein above. Based on the deduced amino acid sequences, we showed that the seven amino acid

residues that were completely conserved in all of the known oleate desaturases, were replaced by different amino acid residues in the only two oleate hydroxylase sequences known at that time.

An object of this invention is a method to convert an oleate hydroxylase to an oleate desaturase. We disclose that an oleate hydroxylase can be converted into an oleate desaturase by changing all seven conserved residues in oleate hydroxylases to the residues that would be found in oleate desaturases. We also show that the same effect can be accomplished by changing six residues. The observation that many combinations of six changes can convert a hydroxylase to a synthetic desaturase shows that no single amino acid change is absolutely required. Thus, there is no amino acid residue that is required for hydroxylase activity but not for desaturase activity. This implies that the functionally significant difference between a hydroxylase and a desaturase is the conformation of the active site as comprised of the conserved amino acid residues. We conclude that changes in the conformation of the active site can change the outcome of the overall reaction.

One implication of this discovery is that any fatty acyl hydroxylase can be converted into a synthetic desaturase by making changes in the conformation of the active site. One aspect of this invention is a procedure for identifying the changes that are made to convert any hydroxylase to a desaturase.

An object of this invention is a method to convert an oleate desaturase into a synthetic oleate hydroxylase by changing all seven conserved residues in oleate desaturases to the residues that would be found in oleate hydroxylases. Based on the analysis of effects of mutations on the amount of desaturase activity exhibited by the LFAH12 hydroxylase, it appears clear that the functionally significant difference between a desaturase and a hydroxylase is the conformation of the active site.

Another implication of this discovery is that any desaturase can be converted into a hydroxylase by making changes in the conformation of the active site. One aspect of this invention is a procedure for identifying the relevant changes that need to be made to convert any desaturase to a synthetic hydroxylase. Changing of a subset of the seven amino acids, as few as four amino acids, results in the conversion of a desaturase to a synthetic hydroxylase. Similarly, changing of the seven, or fewer, amino acids can confer hydroxylase activity on a desaturase.

Thus, these seven amino acid positions define the active site of these fatty acyl enzymes. Besides the genetic techniques exemplified herein, crystallographic and/or spectroscopic techniques may also be used to correlate changes in the active site with enzymatic function. Assays based on determining the chemical and physical properties of the enzymes may be

performed with substrate and/or cofactor analogs to slow or to stabilize the enzymatic reaction. This active site and its function is distinct from the histidine residues previously identified.

The description below applies equally to converting the enzymatic activity from hydroxylase to desaturase, as from desaturase to hydroxylase. In this respect, it should be understood that the functional consequences of modifying a fatty acyl metabolic enzyme (e.g., by genetic engineering) may be assayed by assaying the effects of the modified enzyme on plant fatty acid compounds, especially in seed oil. A modification of the fatty acyl metabolic enzyme may be determined to increase, decrease or not affect any enzymatic activity (e.g., desaturase, hydroxylase) by assaying the fatty acyl content of a cell or plant containing the modified enzyme. A statistically significant increase or decrease in particular desaturated or hydroxylated fatty acids will identify modifications that increase or decrease, respectively, enzymatic activity.

Sequence comparison at the level of amino acid sequence, as well as the functional assays described herein, would show that the number of known nucleotide and amino sequences which are exemplified for fatty desaturases and hydroxylases may be expanded by computer analysis of information found in databases or gathered during sequencing projects to identify related sequences encoding desaturases and hydroxylases. Typically, amino acid sequences are considered to be related with as little as 70% or 80% similarity between the two polypeptides; however, at least 90% or 95% similarity is preferred; and at least 98% similarity is more preferred. Conservative amino acid substitutions may be considered when making sequence comparisons. See generally, Doolittle, *Of URFS and ORFS*, University Science Books, 1986; Gribskov and Devereux, *Sequence Analysis Primer*, Stockton Press, 1991; and references cited therein for algorithms known in the art and used in commercially available software for sequence analysis. A specific example of an algorithm that may be used to calculate sequence divergence is the nucleotide or amino acid versions of the BLAST computer program described by Altschul et al. (J. Mol. Biol., 215, 403-410, 1990; Proc. Natl. Acad. Sci. USA, 87, 5509-5513, 1990; Nucl. Acids Res., 25, 3389-3402, 1997), the complete disclosure of which is fully incorporated herein by reference.

The method according to the present invention is suitable for the production, screening and selection of modified hydroxylase and desaturase enzymes which are derived from naturally existing enzymes. Such mutants are, for example, those encoded by a gene derived from a wild-type FAD2 gene of *A. thaliana* which can be converted to a synthetic oleate hydroxylase. The method can further be advantageously used for the selection of synthetic hydroxylases derived from desaturases other than FAD2-like desaturases. For example, we envision that any

fatty acyl desaturase that shows amino acid sequence similarity to the *A. thaliana* FAD2 gene can be modified according to the teachings of this invention. In particular, it is readily possible to describe the changes necessary to convert the oleate desaturases from soybean, *Zea mays*, or castor to synthetic hydroxylases because the sequences of these enzymes can readily be aligned with the *A. thaliana* FAD2 gene product so that the conserved amino acid residues are aligned (Figure 1). Although the exact numbering of the relevant amino acid residues may change, the intention may be understood by reference to the example in which the numbering of the residues of the LFAH12 and FAD2 gene are aligned. In the LFAH12 gene, positions for substitution of particular interest include 63, 105, 149, 218, 296, 323, 325. The corresponding numbers based on the FAD2 sequence are 63, 104, 148, 217, 295, 322, 324. More generally, we envision that by comparison of the sequences of delta-9 stearoyl-ACP desaturases and stearoyl-9-hydroxylases it will be possible to identify the amino acid residues that are conserved in all delta-9 stearoyl-ACP desaturases but which differ between delta-9 desaturases and 9-hydroxylases. Once such differences have been identified, the knowledge and methods taught herein can be used to create synthetic stearoyl-9-hydroxylases. We also envision that it will be possible to create synthetic hydroxylases for which naturally occurring enzymes are not available.

In addition, a report by.Lee et al. (1997) describes the isolation of a cDNA encoding an acetylenase from *Crepis alpina*. This cDNA is highly similar to plant delta-12-desaturase and the methods used to interconvert the desaturase and hydroxylase functionality may also teach how to interconvert desaturase and acetylenase and vice versa. Additionally, the *Pseudomonas oleovorans* alkane ω-hydroxylase is equally efficient as an epoxidase when presented with 1-octene. Thus the genes encoding fatty acid 12-epoxidases, will also be found in species such as *Euphorbia lagascae* and *Stokesia laevis* that will be closely related to delta-12-desaturase. We envision that these enzymes will be active on linoleate rather than oleate, and will introduce a 12,13 epoxy group. Other lipid enzymes that modify the 12-position such as a ketolase may be related in a similar way as the desaturase and hydroxylase. The methods taught here will also teach how to interconvert any combination of these functionalities.

It will be clear that either oligonucleotide-aided site-directed mutagenesis or region-directed random mutagenesis can be used or any other suitable method for efficiently generating mutations in the hydroxylase or desaturase genes, including complete or partial synthesis of the gene. The method for selecting modified enzymes according to the present invention (which may also include identification, screening, and production) may comprises the following steps: mutagenizing a cloned gene encoding an enzyme of interest or a fragment thereof; isolating the

obtained mutant gene or genes; introducing said mutant gene or genes, preferably on a suitable vector, into a suitable host strain for expression and production; recovering the produced modified enzyme; and identifying those genes encoding modified enzymes having improved properties for application in modifying plant lipid composition. Although the specific examples presented here by way of illustration utilize site-directed mutagenesis, it will be obvious to those skilled in the art that other methods could be used to identify changes that serve equally well for the conversion of a desaturase to a synthetic hydroxylase, or vice versa. Similarly, it will be obvious to those skilled in the art that modified enzymes could also be produced by partial or complete synthesis of modified genes using currently available methods for oligonucleotide synthesis and composition of genes from oligonucleotides and/or fragments from preexisting genes.

Suitable host strains (e.g., bacteria, fungi, yeast, animal cells) for production of enzymes include transformable microorganisms in which expression of the enzymes can be achieved. Specifically, strains of *Saccharomyces cerevisiae* are among the preferred hosts. Expression of fatty acyl enzymes is obtained by using expression signals that function in the selected host organism. Expression signals include sequences of DNA regulating transcription and translation of the fatty acyl metabolizing genes. Proper vectors are able to replicate at sufficiently high copy numbers in the host strain of choice or enable stable maintenance of the introduced gene in the host strain by chromosomal integration.

Assays known in the art may be used to determine and quantify the activity of the modified enzymes in the microbial host. Results provided in the examples show that such results are useful predictors of the activity of the modified enzymes in transgenic plants. The properties of the naturally occurring or mutated enzymes may be enhanced by introducing a variety of mutations in the enzyme. For the most part, the mutations will be substitutions, either conservative or non-conservative, but deletions and insertions may also find use. Another aspect of the invention is the development of novel assays and other processes using the modified enzymes.

For conservative substitutions of "functionally equivalent amino acid residues" the following may be employed for guidance:

Aliphatic neutral non-polar G, A, P, L, I, V

Aliphatic neutral polar C, M, S, T, N, Q

Charged anionic D, E

Cationic K, R

Aromatic F, H, W, Y

where any amino acid may be substituted with any other amino acid in the same chemical category, particularly on the same line. In addition, the polar amino acids N, Q may substitute or be substituted for by the charged amino acids.

The following numbering is based on the *A. thaliana* FAD2 desaturase or, where indicated, the Lesquerella hydroxylase sequence, but the considerations are relevant to other desaturases and hydroxylases having a substantially homologous structure, particularly those having greater than about 70% to 98% similarity. Positions for substitution of particular interest include 63, 105, 149, 218, 296, 323, 325 (numbering based on the LFAH12 sequence). The corresponding positions based on the FAD2 sequence are 63, 104, 148, 217, 295, 322, 324. At some positions there will be an intent to change an amino acid, while maintaining the general conformation and volume of the amino acid at that site.

Substitutions of particular interest include:

63	V or A
105	G or A
149	N or T
218	F or Y
296	V or A
323	A or S
325	I or M

Finally, it will be clear that by deletions or insertions of the amino acids in the desaturase or hydroxylase polypeptide chains, either created artificially by mutagenesis or naturally occurring in desaturases or hydroxylases similar to those described herein, the numbering of the amino acids may change. However, it is to be understood that positions corresponding to amino acid positions of the enzymes descibed herein will fall under the scope of the claims.

Genetic Engineering Applications:

As is well known in the art, the description herein of novel genes encoding plant enzymes that metabolizes fatty acids (i.e., desaturase, hydroxylase, or both activities) allows production of

nucleic acids (e.g., recombinant clones, expression constructs) that could be single- or double-stranded, and comprised of DNA, RNA, modified bases and nucleotides, or combinations thereof. Such polynucleotides may be genomic DNA, cDNA, cRNA, mRNA or heterogeneous RNA (hnRNA). The nucleic acid may contain introns; promoters, enhancers, silencers, transcription initiation/termination sites or other transcriptional regulatory regions; translation initiation/termination sites or other translational regulatory regions; translocation or cellular localization signals; transmembrane regions; regions that regulate message stability; polyadenylation sites; or combinations thereof. For example, a recombinant clone made by genetic engineering may be used to transfect plants, other organisms (e.g., bacteria, fungi, yeast), or cells thereof.

The nucleotide sequences which encode a plant fatty acyl enzyme (e.g., desaturase, hydroxylase, modified versions thereof) may be used in various constructs, for example, as probes to obtain further nucleic acids from the same or other species. Alternatively, these sequences may be used in conjunction with appropriate regulatory sequences to increase levels of the respective fatty acyl enzyme of interest in a host cell for the production of fatty acids with varying amounts of saturation/hydroxylation or study of the enzyme *in vitro* or *in vivo*, or to decrease or increase levels of the respective fatty acyl enzyme of interest for some applications when the host cell is a plant entity, including plant cells, plant parts (including, but not limited to, seeds, cuttings, and tissues), and plants.

A nucleotide sequence encoding a plant fatty acyl enzyme of the present invention may include genomic, cDNA or mRNA derived sequences. By "encoding" is meant that the sequence corresponds to a particular amino acid sequence either in a sense or anti-sense orientation. By "recombinant" is meant that the sequence contains a genetically engineered modification through manipulation via mutagenesis, nucleic acid modifying enzymes, or the like. A cDNA sequence may or may not encode pre-processing sequences, such as transit or signal peptide sequences. Transit or signal peptide sequences facilitate the delivery of the protein to a given organelle and are frequently cleaved from the polypeptide upon entry into the organelle, releasing the "mature" sequence. The use of the precursor DNA sequence is preferred in plant cell expression cassettes.

Furthermore, as discussed above, the complete genomic sequence of a wild-type plant fatty acyl enzyme may be obtained by the screening of a genomic library with a probe, such as a cDNA probe, and isolating those sequences which regulate expression in seed tissue. In this manner, the transcription and translation initiation regions, enhancers, silencers, introns, and/or transcript and translation termination regions of the plant fatty acyl enzyme may be obtained for

use in a variety of nucleic acid constructs, with or without the fatty acyl enyzme structural gene. Thus, nucleotide sequences corresponding to the plant fatty acyl enzyme of the present invention may also provide signal sequences useful to direct transport into an organelle, 5' upstream non-coding regulatory regions (promoters) having useful tissue and timing profiles, 3' downstream non-coding regulatory region useful as transcriptional and/or translational regulatory regions, or may lend insight into other features of the gene.

Once the desired plant fatty acyl enzyme nucleotide sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized. In the structural gene, one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site, or other purposes involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

The nucleotide or amino acid sequences encoding a plant fatty acyl enzyme of the present invention may be combined with other non-native, or "heterologous", sequences in a variety of ways. By "heterologous" sequences is meant any sequence which is not naturally found joined to the plant fatty acyl enzyme, including, for example, combination of nucleotide sequences from the same plant which are not naturally found joined together. Analogously, a "heterologous" nucleic acid describes nucleic acid which is introduced into a host cell or organism which does not naturally contain the nucleic acid.

Using the nucleotide and amino acid sequences disclosed herein, compositions of the present invention may be made substantially pure by overexpressing the nucleic acid or peptide and isolating same. By "substantially pure", a composition of a molecule is described as being at least 80%, preferably at least 90%, more preferably at least 95%, and most preferably at least 99% pure by weight as compared to other substances (i.e., contaminants) of the same chemical character as the recited molecule (e.g., lipid, nucleic acid, protein).

The DNA sequence encoding a plant fatty acyl enzyme of the present invention may be employed in conjunction with all or part of the gene sequences normally associated with the fatty acyl enzyme. In its component parts, a DNA sequence encoding fatty acyl enzyme is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence

encoding plant fatty acyl enzyme, and transcription and translation termination regions.

Potential host cells include both prokaryotic and eukaryotic cells. A host cell may be unicellular or found in a multicellular differentiated or undifferentiated organism depending upon the intended use. Cells of the present invention may be distinguished by having a plant fatty acyl enzyme foreign to the wild-type cell present therein, for example, by having a recombinant nucleic acid construct encoding a plant fatty acyl enzyme therein.

Depending upon the host, the regulatory regions will vary, including regions from viral, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters as well as terminators may be employed. Expression in a microorganism can provide a ready source of the plant enzyme. Among transcriptional initiation regions which have been described are regions from bacterial and fungal (e.g., mold, yeast) hosts, such as *E. coli*, *B. subtilis*, *Saccharomyces cerevisiae*, including promoters such as lacUV5 or a derivative such as trc; bacteriophage T3, T7 or SP6 promoters; trpE; ADC1, Gal1, Gal10, PHO5, or the like.

For the most part, the constructs will involve regulatory regions functional in plants which provide for modified production of plant fatty acyl enzyme with resulting modification of the fatty acid composition. The open reading frame, coding for the plant fatty acyl enzyme or functional fragment thereof will be joined at its 5' end to a transcription initiation regulatory region such as the wild-type sequence naturally found 5' upstream to the fatty acyl enzyme structural gene. Numerous other transcription initiation regions are available which provide for a wide variety of constitutive or regulatable (e.g., inducible) transcription of the structural gene. Among transcriptional initiation regions used for plants are such regions associated with the structural genes such as for nopaline and mannopine synthases, or with napin, soybean β-conglycinin, oleosin, 12S storage protein, the cauliflower mosaic virus 35S promoters, or the like. The transcription/ translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. In embodiments wherein the expression of the fatty acyl metabolic protein is desired in a plant host, the use of all or part of the complete plant fatty acyl enzyme gene is desired, namely all or part of the 5' upstream non-coding regions (promoter) together with the structural gene sequence and 3' downstream non-coding regions may be employed. If a different promoter is desired, such as a promoter native to the plant host of interest or a modified promoter, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source, including the sequence encoding the plant fatty acyl enzyme of interest, or enhanced promoters, such as double

35S CaMV promoters, the sequences may be joined together using standard techniques.

For such applications when 5' upstream non-coding regions are obtained from other genes regulated during seed maturation, those preferentially expressed in plant embryo tissue, such as transcription initiation control regions from the *B. napus* napin gene, or the *Arabidopsis* 12S storage protein, or soybean β-conglycinin (Bray et al., 1987), or the plant fatty acyl hydroxylase promoter are desired. Transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts, are considered desirable for fatty acid modifications in order to minimize any disruptive or adverse effects of the gene product.

Regulatory transcription termination regions may be provided in DNA constructs of the present invention as well. Transcription termination regions may be provided by the DNA sequence encoding the plant fatty acyl enzyme or a convenient transcription termination region derived from a different gene source, for example, the transcription termination region which is naturally associated with the transcription initiation region. Where the transcription termination region is from a different gene source, it will contain at least about 0.5 kb, preferably about 1-3 kb of sequence 3' to the structural gene from which the termination region is derived.

Plant expression or transcription constructs having a plant fatty acyl enzyme as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils. Most especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to rapeseed (canola and high erucic acid varieties), *Crambe*, *Brassica juncea*, *Brassica nigra*, meadowfoam, flax, sunflower, safflower, cotton, *Cuphea*, soybean, peanut, coconut and oil palms and corn. An important criterion in the selection of suitable plants for the introduction on the fatty acyl enzyme is the presence in the host plant of a suitable substrate for the fatty acyl enzyme. Thus, for example, production of vernolic acid will be best accomplished in plants that normally have high levels of linoleic acid in seed lipids.

Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, the present invention is applicable to dicotyledons and monocotyledons species alike and will be readily applicable to new and/or improved transformation and regulation techniques. The method of transformation is not critical to the present invention: various methods of plant transformation are currently available. As newer methods are available to transform crops, they may be directly applied hereunder. For example, many plant species naturally susceptible to *Agrobacterium* infection may be successfully transformed via tripartite or binary vector methods of *Agrobacterium* mediated transformation. In

addition, techniques of microinjection, particle bombardment, and electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

In developing the DNA construct, the various components of the construct or fragments thereof will normally be inserted into a convenient cloning vector which is capable of replication in a bacterial host, e.g., *E. coli*. Numerous vectors exist that have been described in the literature (e.g., plasmid, bacteriophage, cosmid, yeast artificial chromosome or YAC, bacterial artificial chromosome or BAC). After each cloning, the plasmid may be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, insertion, resection, etc., so as to tailor the components of the desired sequence. Once the construct has been completed, it may then be transferred to an appropriate vector for further manipulation in accordance with the manner of transformation of the host cell.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g., antibiotic, heavy metal, toxin, etc., complementation providing prototropy to an auxotrophic host, viral immunity, or the like. Depending upon the number of different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

It is noted that the degeneracy of the DNA code provides that codon substitutions are permissible in the nucleotide sequence contained in nucleic acids of the present invention without any corresponding modification of the amino acid sequence.

As mentioned above, the manner in which the DNA construct is introduced into the plant host is not critical to the present invention. Any method which provides for efficient transformation may be employed. Various methods for plant cell transformation include the use of Ti- or Ri-plasmids, microinjection, electroporation, infiltration, imbibition, DNA particle bombardment, liposome fusion, DNA bombardment, or the like. In many instances, it will be desirable to have the DNA construct bordered on one or both sides of the T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-

DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming plant cells, the expression construct bordered by the T-DNA border(s) will be inserted into a broad host spectrum vector, there being broad host spectrum vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta et al. (1980). Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to kanamycin, the aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to the present invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

Polypeptides with fatty acyl enzymatic activity may be isolated using the identified nucleic acid sequence. The polypeptide may be isolated from natural sources (i.e., plants) or from host cells expressing recombinant fatty acyl enzyme sequences. Polypeptides may be purified using centrifugation, precipitation, specific binding, electrophoresis, and/or chromatography. Separation may be faciliated using enzyme substrates, antibody and/or attachment of a fusion peptide (e.g., avidin, glutathione S-transferase, poly-His, maltose binding protein, myc 9E10-epitope, protein A/G, SV40 T antigen).

Detection of protein expression and localization is facilitated by fusions with reporters such as, for example, alkaline phosphatase (AP), β-galactosidase (LacZ), chloramphenicol acetyltransferase (CAT), β-glucoronidase (GUS), green fluorescent protein (GFP), β-lactamase, luciferase (LUC), or derivatives thereof. Such reporters would use cognate substrates that are preferably assayed by a chromogen, fluorescent, or luminescent signal.

Transcriptional and/or translational fusions of the fatty acyl gene or enzyme and a heterologous nucleic acid or peptide, respectively, may be made. In a transcriptional fusion, a non-translated region of the heterologous gene may be ligated to the fatty acyl metabolic gene or, alternatively, a non-translated region of the fatty acyl gene may be ligated to the heterologous gene. The reading frames of the peptide which is a fatty acyl enzyme and a heterologous peptide may be joined in a translational fusion. If a reporter or selectable marker is used as the heterologous nucleic acid/peptide, then the effect of mutating the nucleotide/amino acid sequences of the fatty acyl enzyme or heterologous nucleic acid/peptide on fatty acid metabolism may be readily assayed. In particular, a transcriptional fusion may be used to localize a regulated promoter of the fatty acyl metabolic gene and a translational fusion may be used to localize the fatty acyl metabolic protein in the cell. For peptide fusions, a peptide recognition site for a protease (e.g., enterokinase, Factor Xa, thrombin) may be included.

Examples.

The following examples which detail materials and methods which can be employed in the practice of the principles of the present invention are offered by way of illustration and not by way of limitation.

Oligonucleotides:

The oligonucleotides used for producing modified forms of the FAD2 gene and the Lfah12 gene are shown in Table 1.

Table 1. Oligonucleotides used for site directed mutagenesis. Letters in uppercase are homologous to the wild type FAD2 or LFAH12 sequences, mutations are in lower case. The mutations made by each oligonucleotide are shown by standard one-letter amino acid abbreviations (e.g., V63A is a mutation that replaces a valine residue at position 63 for an alanine residue).

Name	Use	Amino-acid substitution	Primer sequence (*)
mH1f	a	V63A	ATCACTTTAGcTTCTTGCTTCT
mH1r	a	V63A	AGAAGCAAGAAgCTAAAGTGAT
mH2f	a	G105A	CTGGGTCATTGcCCATGAATGTGGTCACC
mH2r	a	G105A	GGTGACCACATTCATGGgCAATGACCCAG
mH3f	a	N149T	CACCATTCCAACAcTGGATCCCTAGAA
mH3r	a	N149T	TTCTAGGGATCCAgTGTTGGAATGGTG
mH4f	a	F218Y	CATGCACCTATCTaTAAGGACCGTG
mH4r	a	F218Y	CACGGTCCTTAtAGATAGGTGCATG
mH5f	a	V296A	AGAGGAGCTTTGGcTACGGTAGAC
mH5r	a	V296A	GTCTACCGTAgCCAAAGCTCCTCT
mH6f	a	A323S	CATCTCTTTtCAACTATACCGCATT
mH6r	a	A323S	AATGCGGTATAGTTGaAAAGAGATG
mH7F	a	I325m	CATCTCTTTGCAACTATgCCGCATT
mH7r	a	I325M	AATGCGGcATAGTTGCAAAGAGATG
mH67f	a	A323S; I325M	CATCTCTTTtCAACTATgCCGCATT
mH67r	a	A323S; I325M	AATGCGGcATAGTTGaAAAGAGATG
H1f	b	A63V	ATCACTTTAGtTTCTTGCTTCT
H1r	b	A63V	AGAAGCAAGAAaCTAAAGTGAT
H2f	b	A105G	CTGGGTCATTGgCCATGAATGTGGTCACC
H2r	b	A105G	GGTGACCACATTCATGGcCAATGACCCAG
H3f	b	T149N	CACCATTCCAACAaTGGATCCCTAGAA
H3r	b	T149N	TTCTAGGGATCCAtTGTTGGAATGGTG
H4f	b	Y218F	CATGCACCTATCTtTAAGGACCGTG

Table 1. (continued)

Name		Use	Amino-acid substitution	Primer sequence (*)
H4r		b	Y218F	CACGGTCCTTAaaAGATAGGTGCATG
H5f		b	A296V	AGAGGAGCTTTGGtTACGGTAGAC
H5r		b	A296V	GTCTACCGTAaCCAAAGCTCCTCT
H6f		b	S323A	CATCTCTTTgCAACTATACCGCATT
H6r		b	S323A	AATGCGGTATAGTTGcAAAGAGATG
H7f		b	M325I	CATCTCTTTGCAACTATaCCGCATT
H7r		b	M325I	AATGCGGtATAGTTGCAAAGAGATG
mD1f		c	A63V	GACATCATTATAGtCTCATGCTTCTACT
mD1r		c	A63V	AGTAGAAGCATGAGaCTATAATGATGTC
mD2f		c	A104G	CTGGGTCATAGgCCACGAATGCGGTC
mD2r		c	A104G	GACCGCATTCGTGGcCTATGACCCAG
mD3f		c	T148N	CACCATTCCAACAaTGGATCCCTCGAA
mD3r		c	T148N	TTCGAGGGATCCAtTGTTGGAATGGTG
mD4f		c	Y217F	CCCCAACGCTCCCATCTtCAATGACCGAGA
mD4r		c	Y217F	TCTCGGTCATTGaAGATGGGAGCGTTGGGG
mD5f		c	A295V	CAGGGGAGCTTTGGtTACCGTAGACAGAG
mD5r		c	A295V	CTCTGTCTACGGTAaCCAAAGCTCCCCTG
mD67f		c	S322A; M324I	CACCTGTTCgCGACAATaCCGCATTATAACGC
mD67r		c	S322A; M324I	GCGTTATAATGCGGtATTGTCGcGAACAGGTG
H5'	(**)	d	(1)	TATCGAaggcctGATGGGTGCT
H3'	(**)	d	(2)	CTCGCAGTATCgagctCATAACTTATTGTT
D5'	(**)	d	(3)	gatcggtacccgggATGGGTGCAGGTGGAAG-AATGCCGG
D3'	(**)	d	4	gatcgaattcgagctcTCATAACTTATTGTTGTA-CCAGTACACACC

(*) Underlined: target codons for mutagenesis; lower-case letters in bold: oligonucleotide mismatches with the target sequence for the introduction of the described amino-acid substitutions

(**) Lower-case letters in bold: oligonucleotide mismatches with the target sequence for the introduction of restriction sites

(1) Oligonucleotide H5' adds a *Stu*I site immediately before LFAH12 initiating codon.

(2) Oligonucleotide H3' introduces a *Sac*I site following the terminator codon of LFAH12.

(3) Oligonucleotide D5' adds *Kpn*I and *Sma*I sites immediately before FAD2 initiating codon.

(4) Oligonucleotide D3' introduces restriction sites *Sac*I, *Eco*RV, and *Eco*RI sites following the terminator codon of FAD2.

The various uses were (a) introduction of all seven mutations into LFAH12, (b) to revert each of the seven mutations of m7FAH12 to their equivalents in the wt LFAH12 sequence in order to create all combinations of six of the seven changes of m7FAH12, (c) introduction of all seven mutations into FAD2, and (d) introduction of convenient restriction sites in the coding region of genes to facilitate subsequent cloning.

Plasmid Constructions:

The basic construct which was used to create modified hydroxylases from the *A. thaliana* FAD2 desaturase was named pYES-F2. This plasmid was constructed as follows. *A. thaliana* cDNA clone 146M12T7 encoding FAD2 was obtained from the Arabidopsis Stock Center at the Ohio State University. This cDNA sequence was amplified using *Pfu* DNA polymerase (Stratagene) in conjunction with primers designated D5' and D3' to introduce the flanking restriction sites, *Kpn*I and *Sma*I immediately preceding the initiation codon ATG, and *Sac*I and *Eco*RI restriction sites following the terminator codon TGA (Figure 2).

Restriction sites in oligonucleotide D5':

```
5' gatcggtacccgggATGGGTGCAGGTGGAAGAATGCCGG 3'
      ^   ^
   KpnI SmaI
```

Restriction sites in oligonucleotide D3':

```
5' gatcgaattcgagctcTCATAACTTATTGTTGTACCAGTACACACC 3'
      ^     ^
   EcoRI  SacI
```

This amplified wild type FAD2 fragment was cloned into the *Eco*RV site in the vector pZErO (Invitrogen). Following cloning into this high copy bacterial vector, both the coding and noncoding strands of the entire FAD2 insert were sequenced to confirm the presence of the expected sequence and to confirm the absence of secondary mutations that can arise from PCR amplification.

The insert was then excised by restriction with *Kpn*I and *Eco*RI and cloned into the corresponding sites in the bacterial-yeast shuttle vector pYESII (Invitrogen). The pYESII-F2 plasmid was transformed into yeast strain INVSC2 (Invitrogen) by electroporation using a BTX electroporator (BTX).

Plasmid pBNL was used to produce transgenic plants containing the wild type FAD2 gene or modified versions of the gene. This plasmid was constructed as follows. The FAD2 insert in pYES-F2 was excised using the restriction enzymes *Sma*I and *Sac*I. This fragment was cloned into a bacteria-plant shuttle vector pDN, behind the seed specific napin promotor using corresponding restriction sites to produce plasmid pBNL. This construct was introduced into *Agrobacterium tumafaciens* strain GV3101 pMP90 using electroporation (BTX). The agrobacterium was used to transform the FAD2 mutant of *A. thaliana* by vacuum infiltration (Bechtold et al., 1993), and transformants selected for by challenging seeds to germinate on kanamycin containing agarose.

Site-directed Mutagenesis:

Oligonucleotide PCR primers were designed to introduce nucleotide substitutions into LFAH12 and FAD2 through overlap-extension PCR (Ho et al., 1989). These included complementary oligonucleotides 22-32 bp long, encompassing the substitution sites ("mutagenic" primers), and terminal primers. In a first step, overlapping mutagenized fragments were amplified in separate PCR reactions using pairs of mutagenic primers, or a mutagenic primer and a terminal primer for terminal fragments. In a second step, purified overlapping products from the previous steps were assembled and amplified using terminal primers only. Modified LFAH12 genes containing one or seven substitutions were constructed using native-wildtype (WT) coding sequences as templates in the first PCR step.

Modified LFAH12 genes containing only six mutations were constructed using a gene substituted at all seven codons (mLFAH12) as a template. The 5'-end of terminal primers was modified to allow the introduction of convenient restriction sites for the cloning of PCR products. Two sets of PCR "mutagenic" primers were constructed to modify the LFAH12 gene. In the first

set (mH1-5,mH67), the oligonucleotides contained one or two mismatches to modify a target codon(s) in the WT sequence. In the second set (H1-7), the primers contained no mismatch with the WT sequence, and were designed for the substitution of a WT codon for a mutant codon in mLFAH12. In the case of FAD2, only one set of oligonucleotides was synthesized (mD1-5,mD67) which was designed for the modification of the WT gene (Table 1). By genetic engineering, modified desaturase or hydroxylase genes containing each of the seven mutations may be recombined to produce any combination of two, three, four, five or six mutations. As shown below, such modified genes will exhibit varying degrees of metabolic activities.

PCR Conditions:

First step: 10 ng of plasmid DNA was added to a PCR reaction containing 200 µM dNTPs, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM $MgSO_4$, 1% (v/v) Triton X-100, 1000 µg/ml BSA, 3 mM $MgCl_2$, 5% (v/v) DMSO, 125 pmol of each primer, 1.25 units of cloned *Pfu* polymerase (Stratagene), to a final volume of 50 µl. Amplifications conditions were: 4 min denaturation step at 94°C, followed by 30 cycles of 92°C for 1 min, 50°C for 1 min, 72°C for 2 min, concluded with a final extension step at 72°C for 5 min. PCR products were run on and purified from agarose or polyacrylamide gels.

Second step: 10 ng of purified overlapping fragments were used as templates in PCR reactions similar to the above. Amplification conditions were identical except that products were amplified for only 15 cycles.

Cloning Strategy:

PCR fragments encoding modified LFAH12 enzymes were cloned into pBluescriptKS-derived plasmids using one of two general strategies, depending on the location in the molecule of the nucleotides to be substituted. In these approaches, advantage was taken of a unique *Pst*I site near the middle of the coding sequence and that *Pfu* polymerase generates blunt-ended fragments.

If the nucleotide substitutions to be introduced were 5' of the *Pst*I site, overlapping fragments were assembled using terminal primers H5' and mH4r. The resulting products were purified then cut with *Pst*I. In a second step, the pBluescript-derived pLFAH12-1 plasmid (ref) was cut with *Eco*RV and *Pst*I, and the vector fragment was purified and ligated to the cut PCR fragment. If the nucleotide substitutions to be introduced were located 3' of the *Pst*I site, overlapping fragments were assembled using mH3f and H3', cut with *Pst*I, then ligated to the vector fragment from a digest of pLFAH12-1 with *Pst*I and *Sma*I. Alternatively, the assembled

PCR products were cut with *Pst*I and *Sac*I, then ligated to the vector fragment from a restriction digest of pLFAH12-1 with the same enzymes.

Similar strategies were followed to obtain clones containing LFAH12 sequences modified at six out of seven residues, except that mLFAH12 (which encodes a LFAH12 enzyme modified at seven residues) was first substituted for LFAH12 in the vector pLFAH12-1, using the strategies described above, resulting in the vector pmLFAH12. If the nucleotides to be substituted in pmLFAH12 were 5' of the central *Pst*I site, the vector was cut with was *Stu*I and *Pst*I and the insert fragment was substituted for appropriate PCR-assembled fragments cut with *Pst*I. If the nucleotide substitutions to be introduced were 3' of the *Pst*I site, pmLFAH12 was cut with *Pst*I and *Sac*I, the vector fragment was purified, then ligated to appropriate PCR fragments cut with the same enzymes. All inserts were sequenced to confirm the presence of the expected nucleotide substitutions and the absence of secondary mutations which can arise from PCR amplification.

Yeast expression vectors containing WT or modified LFAH12 genes were constructed by excising inserts from the above constructs using the enzymes *Hin*dIII and *Sac*I, and cloning them into the bacterial-yeast shuttle vector pYESII (Invitrogen), cut with the same enzymes. For the construction of binary vectors for plant transformation, pBI121 was cut with *Sma*I and *Sac*I, and the resulting vector fragment was purified. In a second step, inserts were excised from the pBuescriptKS-derived plasmids described above using the enzymes *Stu*I and *Sac*I, and substituted for the GUS gene in the pBI121 vector.

A. thaliana cDNA 146 M12T7 encoding FAD2 was obtained from the Arabidopsis Stock Center at the Ohio State University. This cDNA sequence was amplified using the *Pfu* DNA polymerase in conjunction with primers D5' and D3' to introduce the flanking restriction sites, *Kpn*I and *Sma*I immediately preceding the initiation codon ATG, and *Sac*I and *Eco*RI restriction sites following the terminator codon TGA. This amplified wild type FAD2 fragment was cloned into the *Eco*RV site in the vector pZErO (Invitrogen). For expression of FAD2 in yeast, the insert was then excised by restriction with *Kpn*I and *Eco*RI and cloned into the corresponding sites in the bacterial-yeast shuttle vector pYESII (Invitrogen), resulting in the plasmid pYESII-F2. The binary vector pDN was constructed for seed-specific expression of the WT and modified FAD2 genes in plants. In a first step, the napin promoter was amplified from rapeseed DNA using the oligonucleotide primers nap1 (GGCGTCGACAAGCTTCTGCGGATCAAGCAGCTTTCA) and nap2 (GGTTTTGAGTAGTGATGTCTTGTATGTTCTAGATGGTACCGTAC). In a second step, a *Hin*dIII-*Bgl*II fragment carrying the napin promoter was substituted for the 35S promoter by cutting the pBI121 plasmid (Clontech) with *Hin*dIII and *Bam*HI. A separate excision of the

FAD2 insert was made using the restriction enzymes *Sma*I and *Sac*I. This fragment was cloned into pDN using corresponding restriction sites.

The construction of the mFAD2 cDNA encoding a modified FAD2 enzyme containing seven amino acid substitutions was achieved using overlap extension PCR. Following the second round of assembly-amplification using the primers D5' and D3', the PCR products were treated exactly as the amplified wild type FAD2 sequence described above with respect to the construction of vectors for expression in yeast and in plants. Once again, all inserts were sequenced to confirm the presence of the expected nucleotide substitutions and the absence of secondary mutations which can arise from PCR amplification.

Enzyme Assays:

Root microsomes were prepared as in Miquel and Browse (1992), with some modifications: the extraction buffer contained 2.5 mM NADH and catalase was 10,000 U/ml instead of 2,000 U/ml. After centrifugation at 100,000 g, microsomal membranes were rinsed in desaturase reaction buffer (also containing 10,000 U/ml catalase) before being dispersed in the same buffer to a final concentration of ~0.5 mg/ml microsomal protein. The membranes were then incubated in the presence of 85,000 dpm ^{14}C-oleoyl CoA (52 Ci/mol), and the labeled lipids were extracted after addition to the reaction of an equal volume of 2 M NaCl, 0.2 N HCl and 2 ml of chloroform/methanol (1:1). The chloroform phase was recovered, dried under nitrogen, and the fatty acids were transmethylated in 1 N methanolic HCl for 1 h at 80°C. After addition of an equal volume of 0.9% NaCl, the fatty acid methyl esters (FAME) were extracted into hexane. The hexane was subsequently evaporated under nitrogen, and the FAME were redissolved in 50 ml chloroform. The FAME were then separated along side standards by argentation TLC as in Miquel and Browse (1992), using hexane/ethyl ether (80:20) as the mobile phase. After drying, the plates were exposed to PhosphorImager cassettes, and the radioactivity of target fame's was measured by comparison to known quantities of labeled fatty acids spotted on the plate.

Protein Quantitation:

Proteins were quantitated using Bradford assay reagents (Biorad) or a TCA kit (Pierce) with known amounts of BSA as a standard.

Gene Expression in Yeast:

Yeast strain INVCS2 (InVitrogen) was electroporated with expression constructs and

control vector. Transformed cells were selected on SC-ura plates (obtained from Bio 101) containing 2% glucose. Resulting colonies were used to inoculate SC-ura liquid medium containing 2% galactose. Stationary phase cells were diluted to an OD_{600} of 0.5 in fresh medium, and grown for 5 days at 16°C. The cultures were then centrifuged and pellets were assayed for fatty acid content.

Plant Transformation:

Transgenic plants were generated using a modified *in planta* transformation procedure (Bectold et al., 1993). Batches of 12 to 15 plants were grown on soil covered with nylon screens for 3 to 4 weeks under continuous light (100 mmol $m^{-2}s^{-1}$ irradiation in the 400 to 700 nm range). Primary bolts were removed four days before use to promote growth of multiple secondary bolts. *Agrobacterium tumefaciens* strain GV3101 carrying binary Ti plasmid derivatives was grown in liquid cultures to stationary phase in LB medium with 15 mg/l gentamycin and 50 mg/l kanamycin. Cells were harvested and resuspended in infiltration medium (Murashige and Skoog macro and micronutrient medium containing 10 mg/l 6-benzylaminopurine and 5% glucose). Plants were immersed in the bacterial suspension, then placed under vacuum (600 mm Hg) until tissues appeared uniformly soaked. Infiltrated plants were grown at 25°C under continuous light for four weeks. Seeds, bulk harvested from each pot, were sterilized in a mixture of bleach, water and Triton X-100 (30%, 70%, 0.1%), then germinated on selective medium (1 X Murashige and Skoog salts medium enriched with B5 vitamins, with 50 mg/l kanamycin). Kanamycin resistant seedlings (the T1 generation) were transferred to soil to produce T2 seed.

Measurements of Fatty Acid Composition:

For analysis of seed fatty acid composition, seeds were harvested from dry siliques and the fatty acid composition was determined from lipids extracted from pools of 50 seeds. For analyses of yeast fatty acid composition, cell pellets from 15 ml cultures were extracted in the same way as seed samples.

Fatty acids from samples were transmethylated in 1 ml of 1N methanolic HCl (80°C, 1 hour) and extracted twice into hexane after addition of an equal volume of aqueous 0.9% NaCl. Fatty acid methyl esters were derivatized with BSTFA/TMCS (99:1) at 70°C for 30 min in order to obtain trimethylsilyl fatty acid methyl esters (TMS-FAME) of hydroxylated fatty acids. Fatty acids were resolved on an SP2340 fused silica capillary column (0.25 mm ID, 60 m, Supelco) in splitless mode using 1 ml/min of helium. The injector and detector temperatures were 300°C; the

temperature program was 100 to 160°C at 25°C min, 160 to 240°C at 7°C min, hold at 240°C for 5 min then decrease to 100°C at 25 min. The identity of fatty acids in the samples was determined by comparing retention times and mass spectra to that of standards. A Hewlett-Packard HP5971 MS was used to confirm the identity of eluting compounds.

For the analysis of the composition of individual lipid classes, lipids were extracted as in Miquel and Browse (1992).

Example 1: Modification of a desaturase to a hydroxylase.

Evidence that a desaturase can be converted into a synthetic hydroxylase was obtained by modifying the FAD2 delta-12 oleate desaturase from *A. thaliana* so that it exhibited hydroxylase activity. The nucleotide sequence and corresponding amino acid sequence of the *A. thaliana* FAD2 gene is shown in Figure 3.

A modified version of the FAD2 cDNA encoding seven mutations of the coding sequence and introduction of flanking mutations was achieved using a method based on overlap-extension PCR described by Ho et al. (1989). Briefly, two rounds of PCR are employed (Figure 4). In the first round, in a series of separate PCR reactions, individual fragments designated "A" to "G" are amplified that are tailed by the desired mutated DNA sequence using the wild type *A. thaliana* FAD2 DNA as template and various pairs of oligonucleotide primers encoding the desired DNA mutations. For instance, primers mD1f and mD1r (Table 1) were used in one reaction to amplify a fragment. Similarly, primers mD2f and mD2r (Table 1) are used in a second reaction, mD3f and mD3r are used in a third reaction, mD4f and mD4r are used in a fourth reaction, mD5f and mD5r are used in a fifth reaction and mD67f and mD67r are used in a sixth reaction. The primer pair mD67f and mD67r introduced two amino acid substitutions.

The amplified fragments are then separated from the DNA template by excision and elution from a 5% acrylamide gel. A portion of the purified fragments A-G are pooled into a single PCR reaction and are used as the DNA template in an amplification employing only flanking oligonucleotides D5' and D3' (Table 1). In the first few cycles of amplification, various pairs of DNA fragments overlap and become extended until a continuous template is assembled and subsequently the flanking primers allow for amplification of the entire gene containing all the mutations encoded in the original oligonucleotide primers. It should be emphasized that the fragments are used as a template in the second round of PCR and that full length DNA template is only included in the first round PCR reactions. Following this second round of assembly-

amplification, the fragment is treated exactly as the amplified wild type FAD2 sequence described above with respect to its analysis and introduction into yeast and plants for analysis.

The nucleotide sequence of the mFAD2 gene is presented as SEQ ID NO:1. The amino acid sequence of the polypeptide product of the mFAD2 gene is presented as SEQ ID NO:2. A comparison of the nucleotide sequences of the FAD2 and mFAD2 genes is presented in Figure 5. A comparison of the deduced amino acid squences of the polypeptide products of the FAD2 gene and the mFAD2 gene is presented in Figure 6.

In order to evaluate the effect of the introduced mutations on the activity of FAD2, the modified gene was expressed in yeast. In yeast, FAD2 is active and causes the accumulation of linoleate. Wild-type cells, which do not have this enzymatic activity, do not accumulate this fatty acid. In the present experiment, the mutant desaturase gene (mFAD2) was cloned into the pYESII vector downstream of the GAL1 promoter and electroporated into yeast cells. Transgenic cells were grown under conditions that led to expression of the gene, were harvested and their fatty acid composition determined by gas chromatography. As shown in Table 2, there were dramatic differences between the fatty acid phenotypes of cells expressing the mutant and wild-type genes.

Table 2. Fatty acid composition of wild-type yeast, yeast containing the FAD2 gene and yeast containing the mFAD2 gene The values are average values for five independent transformants. Standard errors are shown in parentheses.

Fatty acid (mol % of total fatty acids)

Line	16:2	18:2	18:1-OH
WT	0.00	0.00	0.00
FAD2	0.58 (0.23)	3.52 (0.57)	0
mFAD2	0.4 (0.17)	1.37 (0.23)	0.51 (0.11)

In cells expressing FAD2, hydroxylated fatty acids were not detectable. However, cells expressing mFAD2 accumulated ricinoleic acid, which constituted on average 0.5% of total fatty acids. Concurrently, their average linoleate content was 1.4% of total fatty acids as compared to 3.5% in cells expressing the unmodified FAD2 gene. Based on the sensitiviy of the assay, we estimate that the effect of the mutations was to increase the ratio of ricinoleate to linoleate content at least 13-fold in transgenic cells. Thus, the seven amino acid differences between FAD2 and mFAD2 convert the enzyme from a desaturase to a synthetic hydroxylase that has both desaturase and hydroxylase activity.

In order to verify that the synthetic hydroxylase was also useful for production of hydroxylated fatty acids in plants, the mutant gene was expressed in transgenic *A. thaliana* plants. In order to detect the lowest possible levels of hydroxylase activity while still measuring oleate desaturation, mFAD2 was expressed in the FAD2 mutant of *A. thaliana* under the control of the strong seed-specific promoter from the *B. rapa* napin gene. We also obtained 15 transgenic lines expressing the wild-type FAD2 desaturase gene. Accumulation of hydroxylated fatty acids was never detectable in the seeds of these plants. Eight transgenic lines expressing mFAD2 were examined (Table 3).

Table 3. Seed fatty acid composition of transgenic *A. thaliana* FAD2 mutant plants expressing the mFAD2 gene. The transgenic plants containing the mFAD2 gene were designated MF2-1 to MF2-8, respectively (average of two measurements on 25 T2 seeds).

	Fatty Acid (mol % of total fatty acids)									
Line	16:0	18:0	18:1	18:2	18:3	20:0	20:1	18:1-OH	18:2-OH	20:1-OH
FAD2*	6	3	40	5	7	1.5	15.2	0	0	0
MF2-1	5.65 (0.07)	3.3 (0.14)	42.1 (2.19)	10.8 (1.83)	12.5 (0.14)	1.4 (0)	20.25 (0.49)	1.35 (0.35)	1.65 (0.35)	0.25 (0.07)
MF2-2	7.25 (0.21)	4.05 (0.21)	28.85 (2.19)	19.8 (1.97)	15.4 (0.84)	1.4 (0)	15.95 (1.62)	2.75 (0.49)	2.4 (0.28)	0.7 (0.14)
MF2-3	5.85 (0.21)	3.2 (0)	34.15 (2.33)	18.3 (1.13)	15.9 (0.28)	1.0 (0)	15.85 (0.49)	2.5 (0.42)	2.85 (0.49)	0.6 (0.14)
MF2-4	6.75 (0.35)	4.7 (0.28)	29.55 (0.49)	23.3 (1.13)	10.55 (0.49)	1.3 (0.14)	13.3 (0.14)	3.45 (0.35)	2.1 (0.42)	0.9 (0.14)
MF2-5	7.65 (0.21)	4.05 (0.21)	27.35 (3.46)	17.95 (1.76)	15.05 (0.49)	1.65 (0.21)	18.25 (1.06)	2.7 (0.56)	3.4 (0.28)	0.7 (0.14)
MF2-6	7.9 (0.42)	4.15 (0.21)	29.3 (2.19)	23.1 (3.53)	13.45 (1.34)	8.8 (0.14)	6.95 (2.75)	3.5 (0.7)	2.4 (0.14)	0.95 (0.07)
MF2-7*	8.3	5	21.7	25	14.1	1.7	13.5	4.9	3	1.3
MF2-8	8.55 (0.35)	4.8 (0.14)	19.45 (2.33)	29.4 (0)	11.9 (0.14)	1.45 (0.07)	12.1 (0.28)	4.75 (0.21)	3.2 (0.28)	1.45 (0.07)

* Replicates were not done for these samples

In contrast with transgenic plants expressing the WT gene, the proportion of hydroxylated fatty acids, which included ricinoleic and derivatives densipolic and lesquerolic acids, ranged from 3.2% to 9.4% (6.7% ± 1.9%) of total seed fatty acids. The ratio of seed linoleate to oleate contents were 2 to 12 times higher (6.4 ± 3.1), which indicated significant desaturase activity, albeit lower than in the seeds of plants transformed with the WT gene. The levels of hydroxylated fatty acid accumulation observed in transgenic plants expressing mFAD2 indicate that all or part of the amino acid substitutions were sufficient to promote significant levels of hydroxylase activity *in planta*. However, these changes did not have the effect of eliminating desaturase

activity of the enzyme.

We envision that one skilled in the art may obtain similar or identical results by practicing minor variations of the invention disclosed herein. One class of modifications is to simply make the corresponding changes in a plant oleate desaturase other than the FAD2 gene from *A. thaliana*. Because of the high degree of sequence conservation among plant microsomal oleate desaturases, identification of other desaturases and their modification by mutagenesis could be performed by the skilled artisan. Another minor variation of this invention would be to omit one or more of the seven amino acid substitutions we have used.

Because of the results disclosed in Example 2, we envision that enzymes with similar or identical properties could be obtained by making only two, three, four, five, or six mutations at the amino acid positions disclosed herein. The minimal set could be identified by systematically making all seven combinations of synthetic hydroxylase enzymes with six out of seven substitutions. In the next step, all of the synthetic hydroxylases with six substitutions that had acceptable levels of hydroxylase activity would be used to design a series of synthetic hydroxylase enzymes with all six combinations of five substitutions and so on until the minimal set of substitutions that gave acceptable activity were identified. We also envision that it may be possible to make synthetic hydroxylase enzymes with similar or identical properties by making more than seven substitutions that included neutral substitutions chosen at random or by comparison of the range of natural variation in desaturases and hydroxylases. We also envision that it may be possible to produce synthetic hydroxylases with similar or identical properties to the enzyme disclosed herein by making different amino acid substitutions at some or all of the seven sites used herein. For example, instead of converting the alanine at position 63 to Valine (mutation A63V), it might be equally effective to convert alanine-63 to isoleucine or leucine. These and other variations on the present invention may be performed by the one of skill in the art.

Example 2: Conversion of a hydroxylase to a desaturase.

In order to increase the ratio of oleate desaturation to oleate hydroxylation catalyzed by the oleate hydroxylase gene from *L. fendleri* (LFAH12), overlap extension PCR with high-fidelity *Pfu* polymerase was used to introduce nucleotide substitutions in the coding region of the LFAH12 gene. As in Example 1, six pairs of mutagenesis primers (designated mH1f to mH67r in

Table 1) were used in combination with terminal primers to amplify fragments which were then assembled in a second PCR amplification step to produce modified full-length coding sequences. The modified gene was designated mFAH12.

The nucleotide sequence of the mFAH12 gene is listed as SEQ ID NO:3. The deduced amino acid sequence of the polypeptide product of the mFAH12 gene is listed as SEQ ID NO:4. A comparison of the nucleotide sequence of the FAH12 and mFAH12 genes is presented as Figure 7. A comparison of the deduced amino acid sequences of the FAH12 and mFAH12 genes is presented as Figure 8.

In order to evaluate the effect of the introduced mutations on the activity of LFAH12, we expressed the modified genes in yeast. In yeast, LFAH12 is active and causes the accumulation of ricinoleate. Wild-type cells do not accumulate ricinoleate. In the present experiment, the mutant hydroxylase gene (mLFAH12) was cloned into the pYESII vector downstream of the GAL1 promoter and electroporated into yeast cells. Induced transgenic cells were harvested and their fatty acid composition determined by gas chromatography. As shown in Table 4, there were dramatic differences between the fatty acid phenotypes of cells expressing the mutant and wild-type genes.

Table 4. Fatty acid composition of wild-type yeast, yeast containing the LFAH12 gene and yeast containing the mLFAH12 gene (average values for five independent transformants).

Fatty Acid (mol% of total fatty acids)

Line	16:2	18:2	18:1-OH
WT	0.00	0.00	0.00
LFAH12	0.74 (0.16)	0.65 (0.03)	1.52 (0.18)
mLFAH12	2.69 (0.27)	6.17 (0.99)	0.33 (0.04)

Although desaturase activity of the LFAH12 enzyme is minor compared to its hydroxylase activity, yeast cells expressing LFAH12 accumulate linoleic and ricinoleic acids to similar levels, possibly because linoleic acid is tolerated better than ricinoleic acid, a phenomenon also observed in plant cells. In cells expressing mFAH12, the ratio of linoleic to ricinoleic acid was on average 43-fold higher than in cells expressing the wild-type gene. There was also a 4-fold increase in the levels of 16:2, which is the product of palmitoleic acid desaturation. These observations suggest a significant increase in desaturase activity associated with a decrease in hydroxylase activity upon introduction of the seven modified amino acid residues in mLFAH12.

To measure the effect of the amino acid substitutions on desaturase activity of the LFAH12 enzyme in plants, mLFAH12 was introduced into the *A. thaliana* FAD2 mutant, which is deficient in oleate desaturation. In transgenic FAD2 plants where LFAH12 is driven by the strong CaMV35S promoter, hydroxylated fatty acids accumulate to high levels, while the mutant phenotype is partially suppressed in roots, due to low levels of desaturase activity of the enzyme. mLFAH12 was expressed under the control of the same promoter and the fatty acid composition of leaves and seeds of transgenic plants was measured.

Table 5. Seed fatty acid composition of transgenic *A. thaliana* FAD2 mutant plants expressing the mFAH12 gene. The transgenic plants containing the mFAH12 gene were designated MH2-1 to MH2-8, respectively (Average of two measurements on 25 T2 seeds)

5

	Fatty Acid (mol% of total fatty acids)					
Line	16:0	16:3	18:0	18:1	18:2	18:3
FAD2	13.75 (0.68)	19.75 (2.12)	0.51 (0.05)	7.78 (1.14)	2.08 (0.33)	44.68 (0.74)
MH2-1*	14.2	13.2	1	9.1	5	49.4
MH2-2*	12.5	15.7	1.8	5.3	10.8	41.4
MH2-3*	13.3	10.8	0.5	3	10.4	58.7
MH2-4	14.2 (0.7)	15.35 (1.06)	0.75 (0.35)	2.85 (0.77)	4.65 (0.49)	55.5 (0.56)
MH2-5	14.35 (0.35)	12.65 (1.76)	0.85 (0.21)	2.15 (1.62)	10.7 (1.69)	58.1 (5.65)
MH2-6	14.05 (0.21)	17.9 (0.84)	1.05 (0.07)	1.0 (0.28)	7.2 (0.42)	51.65 (2.05)
MH2-7	13.35 (0.77)	19.45 (1.2)	0.65 (0.21)	0.9 (0.14)	6.55 (0.21)	58.75 (0.91)
MH2-8	15.85 (0.35)	15.3 (2.54)	1.0 (0)	0.75 (0.35)	8.0 (0.98)	53.6 (0)
MH2-9	18.65 (0.91)	16.8 (1.27)	1.0 (0)	0.5 (0)	5.85 (0.49)	55.65 (0.91)
MH2-10	12.05 (0.77)	19.3 (0.56)	0.5 (0)	0.4 (0.14)	5.35 (0.21)	54.15 (5.72)

In contrast with plants expressing the WT gene which always show a characteristic mutant leaf fatty acid phenotype, expression of mLFAH12 in the FAD2 mutant resulted in suppression of the leaf phenotype in eight out of the 10 transgenic plants which were analyzed in Table 5.

10 Furthermore, analysis of the root fatty acid composition of one of these transgenic lines revealed

that the mutant phenotype was also completely suppressed in this tissue. In addition, hydroxylated fatty acids were not detected in the seeds of any of the transgenic plants. However, expression of mLFH12 resulted in an increase of the ratio of linoleate to oleate content 5 to 10 times (7.6 ± 2.7) over untransformed plants. From this data, which is consistent with the yeast results, we can conclude that expressing mLFAH12 in plants deficient in oleate desaturation has similar phenotypic consequences as expressing a desaturase encoding gene such as FAD2.

We also evaluated the contribution of each of the seven amino acid substitutions to the overall effect of the mutations on the activity of the hydroxylase. We constructed seven modified LFAH12 genes substituted at only six of seven residues, expecting that if the change in enzymatic activity was due for a major part to a single residue, an enzyme with no substitution at this residue would have close to WT activity. Vectors containing each of the seven constructs were introduced into yeast cells, and the accumulation in these cells of ricinoleic and polyunsaturated fatty acids was measured. As shown in Table 6, the fatty acid profiles of cells expressing the different mutant genes was very similar to cells expressing LFAH12.

Table 6. Fatty acid composition of wild-type yeast, yeast containing the LFAH12 gene, yeast containing the mLFAH12 gene and yeast containing mLFAH12 genes substituted at six out of the seven residues. Values given are average values for five independent transformants. Standard errors are given in parentheses.

	Fatty acid		
Line	16:2	18:2	18:1-OH
	mol% of total fatty acids		
WT	0.00	0.00	0.00
LFAH12	0.74 (0.16)	0.65 (0.03)	1.52 (0.18)
mLFAH12	2.69 (0.27)	6.17 (0.99)	0.33 (0.04)
mLF-V63A	3.08 (0.45)	6.98 (0.84)	0.39 (0.04)
mLF-G105A	2.31 (0.17)	5.73 (0.85)	0.47 (0.06)
mLF-N149T	1.65 (0.09)	3.89 (0.48)	0.63 (0.06)
mLF-F218Y	1.85 (0.19)	4.87 (0.48)	0.42 (0.15)
mLF-V296A	1.84 (0.03)	4.22 (0.2)	0.94 (0.03)
mLF-A323S	2.12 (0.16)	4.75 (0.83)	0.39 (0.04)
mLF-I325M	2.76 (0.11)	5.28 (0.36)	0.43 (0.03)

This result indicates that introducing a single WT residue in mLFAH12 is not sufficient to restore the WT activity of the enzyme. In order to eliminate the possibility that more than one residue substitution could alone account for the full effect, we tested the effect of introducing single FAD2 residues in the WT enzyme. From the fatty acid phenotype of transgenic yeast cells expressing the seven mutant genes obtained, none of the mutant enzymes had activities which differed significantly from the WT enzyme.

The above experiment indicates that changing the activity of the *L. fendleri* hydroxylase requires introducing multiple amino acid substitutions in the enzyme. We propose that in the *L. fendleri* hydroxylase, and also in the *A. thaliana* desaturase, a subset of seven residues act together to determine the ratio of desaturase to hydroxylase activities of the enzyme. Because of their proximity to putative iron binding sites, we envision that these residue influence the conformation of the active site. Altering this conformation by introducing key amino acid substitutions would result in affecting the outcome of the overall reaction.

Subsequent experiments to identify amino acid residues that may be responsible for the conversion of the ubiquitous oleate desaturase into an oleate hydroxylase were conducted by again comparing the deduced amino acid sequences of the hydroxylases from *L. fendleri* and *R. communis* with the sequences for oleate desaturases from Arabidopsis, *Zea mays*, *Glycine max* (two sequences) and *R. communis*. Additionally, comparision was also made with the sequence for oleate desaturase from *Brassica napus*. This series of comparisons also revealed that there were only seven residues that were strictly conserved in all of the six desaturases but divergent in both of the available hydroxylases. Four of the residues were adjacent to the conserved histidine clusters. Similar to the initial experiments, the role of these seven residues was assessed by using site-directed mutagenesis to replace the residues found in the *Lesquerella* hydroxylase, LFAH12, with those from the equivalent positions in the desaturases. The seven mutations were V63A, G105A, N149T, F218Y, V296A, A323S, I325M numbered relative to the LFAH12 sequence.

Mutagenic oligonucleotides were used to introduce nucleotide substitutions into cloned genes by overlap-extension PCR (W. Ito, H. Ishiguro, Y. Kurosawa, *Gene* 102, 67 (1991)). In a first step, overlapping fragments were amplified in separate PCR reactions using primer pairs designed to introduce mutations. The products were gel-purified, then assembled in a PCR reaction primed with terminal primers only. Modified LFAH12 genes containing one or seven substitutions were constructed using pLFAH12-1 as template and primers mH1-5 or mH67. Modified LFAH12 genes containing only six mutations were constructed using m_7LFAH12 as a template and one of primers H1-H7 to revert one of the mutations. The 5'-end of terminal primers

was modified to allow the introduction of convenient restriction sites for the cloning of PCR products. The m_7FAD2 was constructed using oligonucleotides D1-5,D67.

The PCR conditions were: 10 ng of plasmid DNA, 200 μM dNTPs, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM $MgSO_4$, 1% (v/v) Triton X-100, 1000 ug/ml BSA, 3 mM $MgCl_2$, 5% (v/v) DMSO, 125 pmol of each primer, 1.25 U of Pfu polymerase (Stratagene), to a final volume of 50 ul. Amplifications conditions were: 4 min denaturation step at 94° C, followed by 30 cycles of 92° C for 1 min, 50° C for 1 min, 72° C for 2 min, concluded with a final extension step at 72° C for 5 min. PCR products were purified from agarose or polyacrylamide gels. For the second PCR step 10 ng of purified overlapping fragments were used as templates in PCR reactions as above except that only 15 cycles were used.

PCR fragments encoding modified LFAH12 enzymes were cloned into pLFAH12-1 cut with *PstI* and one of *SmaI* or *EcoRV* or *SacI*. All inserts were sequenced. Yeast expression vectors containing WT or modified LFAH12 genes were constructed by excising inserts from the above constructs using the enzymes *HindIII* and *SacI*, and cloning them into the *HindIII-SacI* sites of pYESII (Invitrogen). Constructs for plant transformation were made by cloning the *StuI-SacI* fragment from modified LFAH12 genes into the *SmaI-SacI* sites of pBI121.

The FAD2 cDNA clone 146M12T7 was amplified with Pfu DNA polymerase using primers D5' and D3' to introduce restriction sites for *KpnI* and *SmaI* immediately upstream of the initiation codon, and *SacI* and *EcoRI* restriction sites following the terminator codon. The fragment was cloned into the *EcoRV* site in the vector pZErO (Invitrogen). For expression of FAD2 in yeast, the insert was excised by restriction with *KpnI* and *EcoRI* and cloned into the corresponding sites in the pYESII, resulting in plasmid pYESII-F2. Binary Ti-vector pDN was constructed for seed-specific expression of FAD2 genes. In a first step, the napin promoter was amplified from rapeseed DNA using primers ggcgtcgacaagcttctgcggatcaagcagctttca and ggttttgagtagtgatgtcttgtatgttctagatggtaccgtac. A *HindIII-BglII* fragment was cloned into the *HindIII-BglII* sites of pBI121 (Clontech), replacing the 35S promoter. FAD2 coding sequences were excised from pYESII-F2 with *SmaI* and *SacI* and cloned into pDN using corresponding restriction sites.

The construction of the m_7FAD2 cDNA encoding a modified FAD2 enzyme containing seven amino acid substitutions was achieved using overlap extension PCR Following the second round of assembly-amplification using the primers D5' and D3', the PCR products were treated exactly as the amplified wild type FAD2 sequence described above.

Plant expression constructs were introduced into *Agrobacterium tumefaciens* strain

GV3101 pMP90 using electroporation and used to transform Arabidopsis *fad2* mutant plants by vacuum infiltration (D. Bouchez, C. Camilleri, M. Caboche, *Comptes Rendus De L'academie Des Sciences Serie Iii* **316**, 1188 (1993)). The oligonucleotides used were the same as those earlier described Table 1.

In a reciprocal experiment, the seven residues in the Arabidopsis FAD2 oleate desaturase were replaced with the corresponding *Lesquerella* hydroxylase residues. The seven mutations were A63V, A105G, T148N, Y217F, A296V, S322A, M324I based on the numbering of the Arabidopsis FAD2 sequence. The activity of the modified and unmodified genes was then determined by expressing them in yeast and transgenic plants, before analyzing the composition of the total fatty acids. Technical difficulties limited the utility of direct measurements of enzyme activity in cell extracts. The enzymes are integral membrane proteins that act on fatty acids esterified to lipids and require cytochrome b_5 reductase and cytochrome b_5 for activity. The difficulty of quantitatively incorporating labeled lipids into isolated membranes, and ensuring that cytochrome b_5 and b_5 reductase are not limiting, restricts the utility of direct measurements of enzyme activity. Our best estimates of oleate desaturase or oleate hydroxylase activities in crude microsomal preparations from Arabidopsis roots indicated specific activities of 1.2 and 0.3 pmol/mg protein/min, respectively.

The mutant hydroxylase and desaturase genes containing all seven substitutions (designated m_7LFAH12 and m_7FAD2, respectively) were expressed in yeast cells under transcriptional control of the GAL1 promoter. Transgenic cells were harvested after induction and their total fatty acid composition determined by gas chromatography. Wild-type yeast cells do not accumulate detectable levels of diunsaturated or hydroxylated fatty acids (Covello et al. and Kajiwara et al.). The results are shown in Table 7 and presented graphically in Figure 9. Cultures were induced in growth medium containing galactose, ~2x 10^8 cells were harvested, and fatty acids where extracted and modified for analysis by gas chromatography, as described by Broun et al., 1998. Values are the averages (± SE) obtained from five cultures of independent transformants. Expression of FAD2 caused the accumulation of about 4% of diunsaturated fatty acids (16:2 and 18:2) but no detectable hydroxy fatty acids. Expression of LFAH12 caused the accumulation of about 1.4% diunsaturated fatty acids and 1.5% ricinoleic, confirming the mixed function of this enzyme (Broun et al., 1998). Cells expressing m_7FAD2 accumulated ricinoleic acid to ~0.5% of total fatty acids and had ~50% reduction in the accumulation of diunsaturated fatty acids. Thus, replacement of the seven residues converted a strict desaturase to a bifunctional desaturase/hydroxylase comparable in activity to the unmodified *Lesquerella* hydroxylase.

The amount of desaturase activity of the LFAH12 enzyme is relatively low compared to its hydroxylase activity (Broun et al., 1998). However, yeast cells expressing LFAH12 accumulated linoleic and ricinoleic acids to similar levels, possibly because linoleic acid is more stable than ricinoleic acid in yeast cells. In cells expressing m_7FAH12, the ratio of 18:2 to ricinoleic acid was on average 43-fold higher than in cells expressing LFAH12. There was also a 16-fold increase in the ratio of 16:2 to ricinoleic acid. Thus, there was both a major increase in desaturase activity and a decrease in hydroxylase activity upon introduction of the seven desaturase-equivalent residues into LFAH12.

Table 7

% Total Fatty Acids/Fatty Acid

	Line	16:2	18:2	R*
gene	FAD2	0.58	3.52	0.00
	M7FA22	0.40	1.37	0.51
	LFAH12	0.74	0.65	1.52
	m7LFAH12	2.69	6.17	0.33

*Ricinoleate

The activity of the mutant enzymes *in planta* was examined by using the corresponding genes to produce stable transgenic plants in an Arabidopsis *fad2* mutant which is deficient in oleate desaturase activity (Miguel et al., 1992). The results are shown in Table 8 and presented graphically in Figure 10. Measurements were made of the fatty acid composition of leaf lipids from wild type, the *fad2* mutant, and transgenic *fad2* plants expressing LFAH12 or m_7LFAH12, under the control of the CAMV 35S promoter. Values at means ± SE (n=3). Expression of LFAH12 under transcriptional control of the constitutive CAMV 35S promoter resulted in accumulation of high levels of hydroxy fatty acids in seeds, but no detectable suppression of the *fad2* mutant phenotype in leaves. By contrast, expression of m_7LFAH12 under the same circumstances resulted in complete suppression of the *fad2* phenotype in 8 out of 10 transgenic plants analyzed. There was an average 21-fold increase in the ratio of linoleate to oleate in leaf fatty acids and a small increase in the amount of linolenic acid. These results, which are consistent with the results of the yeast assays, confirm that expression of m_7LFAH12 in plants deficient in oleate desaturation has identical phenotypic consequences to expressing a wild type desaturase such as FAD2 (Miguel et al., 1992).

Table 8

% total Fatty Acids/gene

		Wild Type	fad2	fad2(LFAH12)	fad2(m7LFAH12)
Fatty Acid	16:0	12.57	11.40	10.93	12.14
	16:3	13.20	16.45	17.00	13.89
	18:0	1.0	0.75	0.50	1.20
	18:1	3.13	19.75	20.30	3.10
	18:2	14.33	3.95	3.70	13.09
	18:3	47.37	38.25	39.77	47.37

To evaluate the effect of the seven mutations on the activity of the FAD2 gene, FAD2 and m_7FAD2 were expressed in the Arabidopsis *fad2* mutant under the control of the strong seed-specific promoter from the *B. rapa* napin gene (Miguel et al., 1992). The results are shown in Table 9 and presented graphically in Figure 11. The abbreviations used in Figure 11 are: ricinoleic acid (18:1-OH), densipolic acid (18:2-OH) and lesquerolic acid (20:1-OH).

As expected from previous studies (Broun et al., 1998), none of the 15 transgenic lines expressing the FAD2 gene accumulated detectable hydroxy fatty acids, although the ratio of linoleate to oleate accumulation was increased an average of 10-fold as compared to untransformed controls. In the transgenic lines expressing m_7FAD2, the amount of hydroxylated fatty acids, which included ricinoleic, densipolic and lesquerolic acids, comprised up to 9.4% of total seed fatty acids. The ratio of seed linoleate to oleate contents was increased an average of 6.4-fold (results not presented), which indicated that m_7FAD2 exhibited significant desaturase activity, albeit lower than in the seeds of plants transformed with the wild type FAD2 gene. The high levels of hydroxy fatty acid accumulation observed in transgenic plants expressing m_7FAD2 indicated that the modified desaturase had comparable levels of hydroxylase activity, in the *in planta* assay, to the native *Lesquerella* hydroxylase enzyme. However, the seven amino acid substitutions did not completely eliminate the desaturase activity of the enzyme.

Table 9

% Total fatty acid/Fatty Acid

		20:1-OH	18:2-OH	18:1-OH
T.L.*	1	2.70	3.40	0.70
	2	3.50	2.40	0.95
	3	4.90	3.00	1.30
	4	4.75	3.20	1.45
	5	1.72	1.82	2.27
	6	2.98	2.39	1.55
	7	2.76	3.86	1.88
	8	3.14	4.50	1.18

* Transgenic Line

Two approaches were used to determine whether any single amino acid residue of the seven had a major effect on the ratio of hydroxylase to desaturase activities. First, each of the seven FAD2-equivalent residues were individually introduced into the LFAH12 enzyme. None of the enzymes containing single amino acid substitutions had activities that differed significantly from the wild type hydroxylase enzyme when expressed in yeast. We also constructed seven modified LFAH12 genes containing all combinations of six desaturase-equivalent residues. The seven constructs were introduced into yeast cells, and the accumulation in these cells of ricinoleic and polyunsaturated fatty acids was measured. The results are shown in Table 10 and are shown graphically in Figure 12. Seven derivatives of the m_7LFAH12 gene containing all combinations of six out of seven substitutions were introduced into yeast cells and the fatty acid composition of five independent cultures was measured. The "X" designation refers to the unmodified amino acid (i.e., enzyme XI325M contains all of the seven substitutions *except* I325M). Each of the seven lines exhibited a ratio of diunsaturated/hydroxylated fatty acids that was closer to the ratio produced by the m_7FAH12 enzyme than by FAH12. Thus, for the *Lesquerella* hydroxylase, and presumably also for the Arabidopsis desaturase, as few as six residues principally determine the ratio of the functional outcome in terms of desaturation or hydroxylation of the enzyme. All lines showed somewhat reduced levels of desaturase activity, with the largest reductions of ~40% seen in F218Y and G105A. Therefore, we made a construct in which both these changes were combined (xF218Y/G105A). This construct exhibited similar activity to the individual F218Y and G105A mutants suggesting that their effects are redundant and that the observed changes in activity result from interactions of more than two of the seven residues. Considered together, these results indicate that no single amino acid position plays an essential role in catalytic outcome.

Rather, changes in activity result from a combined effect of several amino acid positions which have partially overlapping effects.

Table 10

% Total Fatty Acids/Fatty Acid

		18:2-OH	18:2	16:2
A.A.S.	X1325M	0.33	6.17	2.69
	XA323S	0.43	5.28	2.76
	XV296A	0.39	4.75	2.12
	XF218Y	0.94	4.22	1.84
	XN149T	0.42	4.87	1.85
	XG105A	0.63	3.89	1.65
	XV63A	0.47	5.73	2.31
	M7LFAH12	0.39	6.98	3.08
	LFAH12	1.52	0.65	0.74

*Amino Acid Substitutions

Because four of the seven amino acids are adjacent to histidine residues that have been identified as essential to catalysis, we hypothesized that these four residues may be of greatest importance to the outcome of the reaction. A modified FAD2 enzyme, designated m_4FAD2, was constructed in which these four amino acids were replaced by their equivalents from the Lesquerella hydroxylase A104G, T148N, S322A, M324I. Expression of m_4FAD2 in seeds of wild type Arabidopsis resulted in the accumulation of average levels of hydroxy fatty acids that were similar to those obtained with m_7FAD2 (Fig. 11). Thus, only four changes are required to convert a strict desaturase to an enzyme which retains some desaturase activity but is also an efficient hydroxylase.

Biochemical and structural similarities between the desaturase and hydroxylase in addition to recent kinetic isotope experiments, suggest that there is an initial oxidation event at C-12 for both enzymes (Buist et al., 1998). We envision that since no specific single amino acid change is required, and in view of the substantial effect of the four residues that about the active site histidines, that the differences between desaturase and hydroxylase outcome is influenced by the geometry of the active site. The differences likely reflect changes in the relative positioning of the substrate with respect to an activated oxygen species, such that the conformation of the m_4(or m_7)FAD2, or wild type LFAH12 favors oxygen transfer rather than a second C-H bond cleavage

at C-13. This mode of evolving new catalytic activity departs from the accepted paradigm in which the evolution of new activities "involves the incorporation of new catalytic groups into the active site" (Babbitt et al., 1997).

Previous studies have shown how site specific mutagenesis can alter the specificity of enzymes, both for substrates and in terms of regiospecificity (Yuan et al., 1995; Sloane et al., 1991; and Cahoon et al., 1997). The functional outcome of an enzymatic reaction has also been altered from oxidase to oxygenase for the F208Y mutant of ribonucleotide reductase, but this was capable only of single turnover resulting in the formation of dopa-208. In contrast, the experiments described here demonstrate that a desaturase can be engineered to perform efficient hydroxylation by as few as four amino acid changes. And, conversely for a hydroxylase, the ratio of desaturation to hydroxylation can be greatly changed in favor of desaturation by changing as few as six residues. The resulting enzymes are catalytically active *in vivo* and their expression in transgenic Arabidopsis results in the accumulation of substantial levels of modified fatty acids.

The results presented here provide an insight into catalytic flexibility of diiron-containing enzymes. In addition to desaturases and hydroxylases, Stymne and collaborators have recently discovered that acetylenic and epoxy fatty acids are produced by desaturation and epoxidation of double bonds by enzymes that are structurally similar to the enzymes described here (Lee et al. 1998). Thus, it appears that variations of the same catalytic center can catalyze the formation of at least four different functional groups. Since various combinations of these four functional groups define most of the chemical complexity found among the hundreds of different fatty acids that occur in higher plants, it is now apparent that most of the chemical complexity of plant fatty acids can be accounted for by divergence of a small number of desaturases. Extrapolating from the results described here, it also seems very likely that a small number of amino acid substitutions will account for the functional divergence of desaturases, hydroxylases, expoxgenases, and acetylenic-bond forming enzymes.

Although the present invention has been described in detail with reference to its presently preferred embodiments, it will be understood by those of ordinary skill in the art that various modifications and improvements to the present invention are believed to be apparent to one skilled in the art. The embodiments were chosen and described in order to explain the principles of the invention and its practical application to thereby enable others skilled in the art to best utilize the invention in various embodiments and with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the claims appended hereto.

LITERATURE CITED

Babbitt, P.C. and J.A. Gerlt, *J. Biol. Chem.* 272, 30591 (1997).

Bechtold, N., Ellis, J., Pelletier, G. (1993) In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C.R. Acad. Sci. Paris 316, 1194-1199.

Broun, P., S. Boddupalli, C. R. Somerville, *Plant J.* 13, 201 (1998), A bifunctional oleate 12-hydroxylase: Desaturase from *Lesquerella fendleri.*

Broun, P., Somerville, C.R. (1997) Accumulation of ricinoleic, lesquerolic, and densipolic acids in seeds of transgenic *A. thaliana* plants that express a fatty acyl hydroxylase cDNA from castor bean. Plant Physiol. 113, 933-942.

Buist, P.H. and B. Behrouzian *J. Am. Chem. Soc.* 120, 871 (1998)

Cahoon, E.B., Y. Lindqvist, G. Schneider, J. Shanklin, *Proc. Natl. Acad. Sci. USA* 94, 4872 (1997).

Covello, P.S. and D.W. Reed, *Plant Physiol.* 111, 223 (1996)

Frey, M. et al., *Science* 277,696 (1997)

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51-9.

Kajiwara, S. et al., *J. Bacteriol.* 62,4309 (1996)

Kim, C., Y. Dong, L. Que, *J. Am. Chem. Soc.* 119, 3635 (1997)

Kridl, J.C. et al., *Seed Sci. Res.* 1, 209 (1991).

Lee, M. et al., *Science* 280, 915 (1998).

Lee, M., Lenman, M., Banas, A., Bafor, M., Sjodahl, S., Dahlqvist A., Gummeson, P., Schweizer, M., Singh, S., Nilsson, R., Liljenberg, C., Stymne, S. (1997) Cloning of a cDNA encoding an acetylenic acid forming enzyme and its functional expression in yeast. Abstract of the Biochemistry and Molecular Biology of Plant Fatty Acids and Glycerolipids Symposium, Lake Tahoe, CA, June 4-8 1997, abstract A7.

Lindqvist, Y., W. Huang, G. Schneider, J. Shanklin, *EMBO J.* 15, 4081 (1996).

Miquel, M., Browse, J. (1992) Arabidopsis mutants deficient in polyunsaturated fatty acid synthesis. J. Biol. Chem. 267, 1502-1509.

Moreau, R.A. and P.K. Stumpf *Plant Physiol.* 67, 672 (1981)

Pikus, J.D. et al., *Biochemistry* 35, 9106 (1996).

Que, L., *Science* 253, 273 (1991).

Que, L. and Y. Dong, *Acc. Chem. Res.* 29, 190 (1996)

Shanklin, J., Achim, C., Schmidt, H., Fox, B.G., Munck, E. (1997) Mossbauer studies of alkane omega-hydroxylase: Evidence for a diiron cluster in an integral-membrane enzyme. Proc. Natl. Acad. Sci. USA 94, 2981-2986.

Shanklin, J. and E.B. Cahoon, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 49, 611 (1998).

Shanklin, J. and C.R. Somerville, *Proc. Natl. Acad. Sci. USA* 88, 2510 (1991).

Shu, L., J.C. Jesheim, K. Kauffmann, E. Münck, J.D. Lipscomb, L. Que, *Science* 275, 515 (1997).

Sloane, D.L., R. Leung, C.S. Craik, E. Sigal, *Nature* 354, 149 (1991)

Suzuki, M., T. Hayakawa, J. P. Shaw, M. Rekik, S. Harayama, *J Bacteriol* 173, 1690 (1991).

van de Loo, F., Broun, P., Turner, S., Somerville, C. (1995) An oleate 12-hydroxylase from castor (*Ricinus communis*) is a fatty acyl desaturase homolog. Proc. Natl. Acad. Sci. USA 92, 6743-6747.

van de Loo, F.N., B. Fox, C. R. Somerville, in *Plant Lipids* T. Moore, Ed. (CRC Press, Boca Raton, 1993) pp. 91-126.

Wallar; B.J. and J.D. Lipscomb, *Chem. Rev.* **96**, 2625 (1996)

Wang, J. and E. Pichersky *Arch. Biochem. Biophys.* **349**,153 (1998).

Yuan, L., T.A. Voelker, D.J. Hawkins, *Proc. Natl. Acad. Sci. USA* **92**, 10639 (1995)

We Claim:

1. A mutant fatty acyl desaturase which has fatty acyl hydroxylase activity.

2. A modified oleate desaturase in which at least two amino acid substitutions has been made to a native oleate desaturase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

3. A modified oleate desaturase in which at least four amino acid substitutions has been made to a native oleate desaturase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

4. A modified oleate desaturase in which at least six amino acid substitutions has been made to a native oleate desaturase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

5. A modified oleate desaturase, which has been aligned for maximal amino acid sequence similarity with FAD2 oleate desaturase, in which the amino acid sequence is numbered to correspond to the numbering of the FAD2 oleate desaturase, and in which the following amino acid substitutions have been made A63V, A104G, T148N, Y217F, A295V, S322A, and M324I.

6. A modified fatty acyl desaturase in which at least the active site of a native desaturase has been mutated such that an enzymatic activity of the native desaturase has been altered but specificity for fatty acyl substrate is retained.

7. A transgenic plant containing a modified desaturase gene that has been modified so as to catalyze hydroxylation of a fatty acyl substrate of the non-modified desaturase gene.

8. A transgenic plant containing a modified oleate desaturase gene of any one of claims 2-5 that has been modified so as to catalyze hydroxylation of oleate.

9. A transgenic plant containing a gene encoding the modified desaturase of any one

of claims 1-6.

10. Oil or other fatty acyl compounds produced by a modified or mutant desaturase.

11. A method of modifying a fatty acyl desaturase to a fatty acyl hydroxylase consisting of identifying and changing amino acid residues that are conserved in functionally equivalent desaturase enzymes from various plant species but that are not identical in fatty acyl hydroxylases that exhibit significant overall sequence similarity to the fatty acyl desaturases, and which catalyze hydroxylation at one of the carbon residues on the fatty acyl substrate that is desaturated by the corresponding desaturase; said modifications being made by changing the amino acid residue so that it is identical or functionally equivalent to the amino acid residue found in the naturally occurring hydroxylase.

12. A mutant fatty acyl hydroxylase which has fatty acyl desaturase activity.

13. A modified oleate hydroxylase in which at least two amino acid substitutions has been made to a native oleate hydroxylase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

14. A modified oleate hydroxylase in which at least four amino acid substitutions has been made to a native oleate hydroxylase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

15. A modified oleate hydroxylase in which at least six amino acid substitutions has been made to a native oleate hydroxylase at an amino acid position selected from the group consisting of 69, 111, 155, 226, 304, 331, and 333 as numbered in Figure 1.

16. A modified oleate hydroxylase, which has been aligned for maximal amino acid sequence similarity with FAD2 oleate hydroxylase, in which the amino acid sequence is numbered to correspond to the numbering of the FAD2 oleate hydroxylase, and in which the following amino acid substitutions have been made V63A, G105A, N149T, F218Y, V296A, A323S, and I325M.

17. A modified fatty acyl hydroxylase in which at least the active site of a native hydroxylase has been mutated such that an enzymatic activity of the native hydroxylase has been altered but specificity for fatty acyl substrate is retained.

18. A transgenic plant containing a modified hydroxylase gene that has been modified so as to catalyze desaturation of a fatty acyl substrate of the non-modified hydroxylase gene.

19. A transgenic plant containing a modified oleate hydroxylase gene of any one of claims 13-16 that has been modified so as to catalyze desaturation of oleate.

20. A transgenic plant containing a gene encoding the modified hydroxylase of any one of claims 12-17.

21. Oil or other fatty acyl compounds produced by a modified or mutant hydroxylase.

22. A method of modifying a fatty acyl hydroxylase to a fatty acyl desaturase consisting of identifying and changing amino acid residues that are conserved in functionally equivalent desaturase enzymes from various plant species but that are not identical in fatty acyl hydroxylases that exhibit significant overall sequence similarity to the fatty acyl desaturases, and which catalyze hydroxylation at one of the carbon residues on the fatty acyl substrate that is desaturated by the corresponding desaturase; said modifications being made by changing the amino acid residue so that it is identical or functionally equivalent to the amino acid residue found in the naturally occurring desaturase.

23. A method for altering fatty acid-modifying enzymes such as desaturases, hydroxylases, epoxidases and acetylene-forming enzymes so that the product or products of the reaction catalyzed by the modified enzyme more closely resemble those produced by a different functional class of enzymes than the unmodified enzyme comprising:

a) identifying a number of amino acid sequences for each of two classes of functionally distinguishable but structurally related enzymes,

b) aligning the sequences for maximal sequence identity or similarity,

c) identifying those critical amino acid residues which are identical in all members of one functional class of enzymes but differ in the other class of enzymes,

d) altering the gene or genes encoding one class of enzymes so that the critical amino acid residues of the modified enzyme are changed to more closely resemble those found at the corresponding positions in the other class of enzymes, and

e) obtaining expression of the modified gene or genes in a host organism that is capable of transcribing and translating the gene to produce the modified enzymes of interest.

```
                        10         20         30         40         50
LFFAH12      1 MGAGGRIM-- --VTPSSKKS --ETEALKRG PCEKPPFTVK DLKKAIPQHC   50
FAH12        1 MGGGGRMSTV ITSNNSEKKG --GSSHLKRA PHTKPPFTLG DLKRAIPPHC   50
ATFAD2       1 MGAGGRMP-- --VPTSSKKS --ETDTTKRV PCEKPPFSVG DLKKAIPPHC   50
GMFAD2-1     1 MGLA-KETTM GGRGRVAKVE VQGKKPLSRV PNTKPPFTVG QLKKAIPPHC   50
GMFAD2-2     1 MGAGGR---- TDVPPANRKS --EVDPLKRV PFEKPQFSLS QIKKAIPPHC   50
ZMFAD2       1 MGAGGRMTEK EREKQEQLAR ATGGAAMQRS PVEKPPFTLG QIKKAIPPHC   50

                        60         70         80         90        100
LFFAH12     51 FKRSIPRSFS YLLTDITLVS CFYYVATNYF SLLPQPLSTY LAWPLYWVCQ   100
FAH12       51 FERSFVRSFS YVAYDVCLSF LFYSIATNFF PYISSPLS-Y VAWLVYWLFQ   100
ATFAD2      51 FKRSIPRSFS YLISDIIIAS CFYYVATNYF SLLPQPLS-Y LAWPLYWACQ   100
GMFAD2-1    51 FQRSLLTSFS YVVYDLSFAF IFY-IATTYF HLLPQPFS-L IAWPIYWVLQ   100
GMFAD2-2    51 FQRSVLRSFS YVVYDLTIAF CLYYVATHYF HLLPGPLS-F RGMAIYWAVQ   100
ZMFAD2      51 FERSVLKSFS YVVHDLVIAA ALLYFALAII PALPSPLR-Y AAWPLYWIAQ   100

                       110        120        130        140        150
LFFAH12    101 GCVLTGIWVI GHECGHHAFS DYQWVDDTVG FIFHSFLLVP YFSWKYSHRR   150
FAH12      101 GCILTGLWVI GHECGHHAFS EYQLADDIVG LIVHSALLVP YFSWKYSHRR   150
ATFAD2     101 GCVLTGIWVI AHECGHHAFS DYQWLDDTVG LIFHSFLLVP YFSWKYSHRR   150
GMFAD2-1   101 GCLLTGVWVI AHECGHHAFS KYQWVDDVVG LTLHSTLLVP YFSWKISHRR   150
GMFAD2-2   101 GCILTGVWVI AHECGHHAFS DYQLLDDIVG LILHSALLVP YFSWKYSHRR   150
ZMFAD2     101 G---------- -------AFS DYSLLDDVVG LVLHSSLMVP YFSWKYSHRR   150

                       160        170        180        190        200
LFFAH12    151 HHSNNGSLEK DEVFVPPKKA AVKWYVKYL- NNPLGRILVL TVQFILGWPL   200
FAH12      151 HHSNIGSLER DEVFVPKSKS KISWYSKYS- NNPPGRVLTL AATLLLGWPL   200
ATFAD2     151 HHSNTGSLER DEVFVPKQKS AIKWYGKYL- NNPLGRIMML TVQFVLGWPL   200
GMFAD2-1   151 HHSNTGSLDR DEVFVPKPKS KVAWFSKYL- NNPLGRAVSL LVTLTIGWPM   200
GMFAD2-2   151 HHSNTGSLER DEVFVPKQKS CIKWYSKYL- NNPPGRVLTL AVTLTLGWPL   200
ZMFAD2     151 HHSNTGSLER DEVFVPKKKE ALPWYTPYVY NNPVGRVVHI VVQLTLGWPL   200

                       210        220        230        240        250
LFFAH12    201 YLAFNVSGRP YDG-FASHFF PHAPIFKDRE RLQIYISDAG ILAVCYGLYR   250
FAH12      201 YLAFNVSGRP YDR-FACHYD PYGPIFSERE RLQIYIADLG IFATTFVLYQ   250
ATFAD2     201 YLAFNVSGRP YDG-FACHFF PNAPIYNDRE RLQIYLSDAG ILAVCFGLYR   250
GMFAD2-1   201 YLAFNVSGRP YDS-FASHYH PYAPIYSNRE RLLIYVSDVA LFSVTYSLYR   250
GMFAD2-2   201 YLALNVSGRP YDR-FACHYD PYGPIYSDRE RLQIYISDAG VLAVVYGLFR   250
ZMFAD2     201 YLATNASGRP YPR-FACHFD PYGPIYNDRE RAQIFVSDAG VVAVAFGLYK   250

                       260        270        280        290        300
LFFAH12    251 YAASQGLTAM ICVYGVPLLI VNFFLVLVTF LQHTHPSLPH YDSTEWEWIR   300
FAH12      251 ATMAKGLAWV MRIYGVPLLI VNCFLVMITY LQHTHPAIPR YGSSEWDWLR   300
ATFAD2     251 YAAAQGMASM ICLYGVPLLI VNAFLVLITY LQHTHPSLPH YDSSEWDWLR   300
GMFAD2-1   251 VATLKGLVWL LCVYGVPLLI VNGFLVTITY LQHTHFALPH YDSSEWDWLK   300
GMFAD2-2   251 LAMAKGLAWV VCVYGVPLLV VNGFLVLITF LQHTHPALPH YTSSEWDWLR   300
ZMFAD2     251 LAAAFGVWWV VRVYAVPLLI VNAWLVLITY LQHTHPSLPH YDSSEWDWLR   300

                       310        320        330        340        350
LFFAH12    301 GALVTVDRDY GILNKVFHNI TDTHVAHHLF ATIPHYNAME ATEAIKPILG   350
FAH12      301 GAMVTVDRDY GVLNKVFHNI ADTHVAHHLF ATVPHYHAME ATKAIKPIMG   350
ATFAD2     301 GALATVDRDY GILNKVFHNI TDTHVAHHLF STMPHYNAME ATKAIKPILG   350
GMFAD2-1   301 GALATMDRDY GILNKVFHHI TDTHVAHHLF STMPHYHAME ATNAIKPILG   350
GMFAD2-2   301 GALATVDRDY GILNKVFHNI TDTHVAHHLF STMPHYHAME ATKAIKPILG   350
ZMFAD2     301 GALATMDRDY GILNRVFHNI TDTHVAHHLF STMPHYHAME ATKAIRPILG   350

                       360        370        380        390        400
LFFAH12    351 DYYHFDGTPW YVAMYREAKE CLYVEPDTER GKKGVYYYNN K-L.......   400
FAH12      351 EYYRYDGTPF YKALWREAKE CLFVEPDEGA PTQGVFWYRN KY-.......   400
ATFAD2     351 DYYQFDGTPW YVAMYREAKE CIYVEPDREG DKKGVYWYNN K-L.......   400
GMFAD2-1   351 EYYQFDDTPF YKALWREARE CLYVEPDEGT SEKGVYWYRN KY-.......   400
GMFAD2-2   351 EYYRFDETPF VKAMWREARE CIYVEPDQST ESKGVFWYNN KL-.......   400
ZMFAD2     351 DYYHFDPTPV AKATWREAGE CIYVEPE--- DRKGVFWYNK KF*.......   400
```

Fig1

General strategy for site directed mutagenesis of Arabidopsis *Fad2*

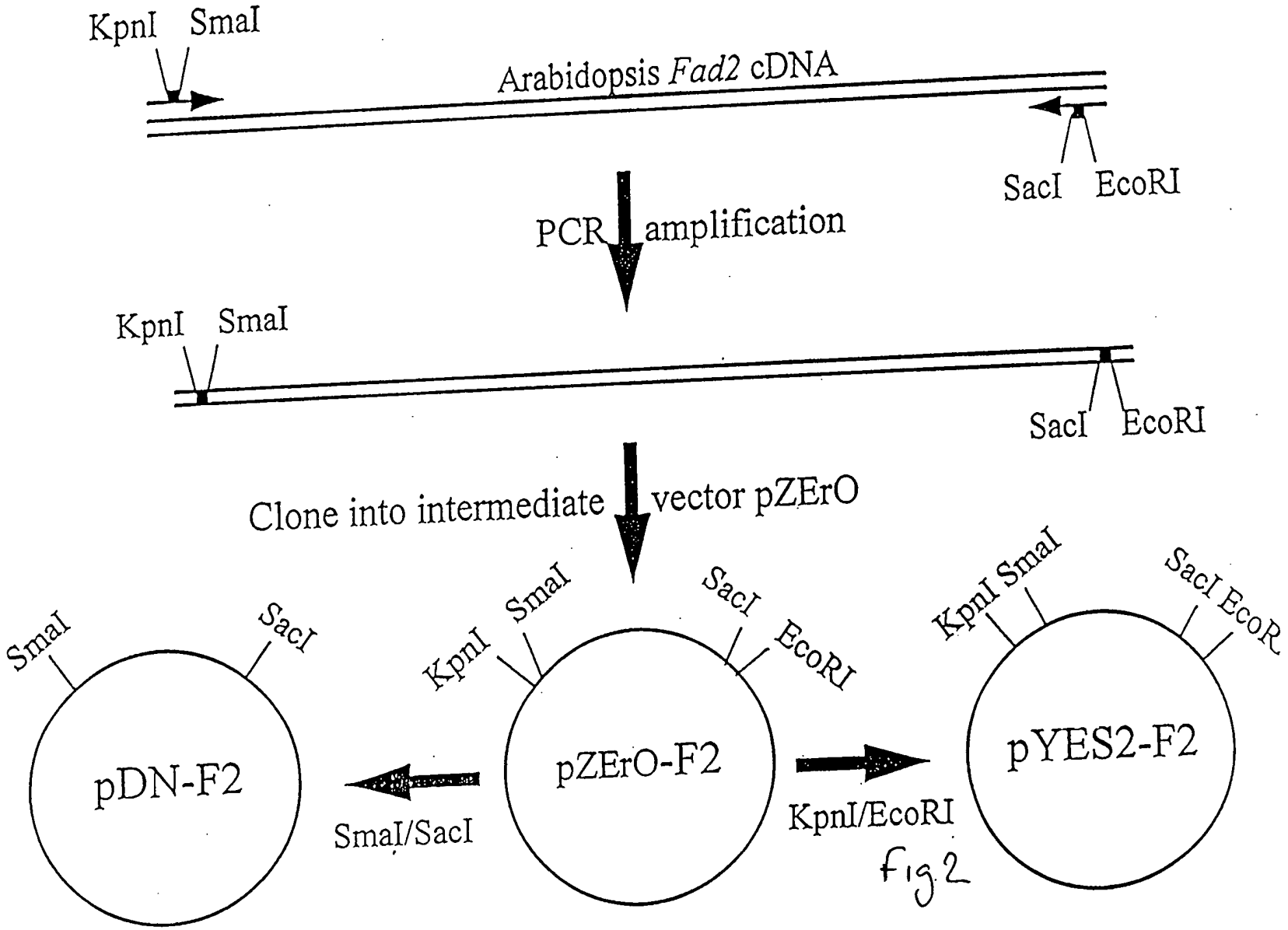

fig 2

```
  1  Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser   16
  1  ATG GGT GCA GGT GGA AGA ATG CCG GTT CCT ACT TCT TCC AAG AAA TCG   48

 17  Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser   32
 49  GAA ACC GAC ACC ACA AAG CGT GTG CCG TGC GAG AAA CCG CCT TTC TCG   96

 33  Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser   48
 97  GTG GGA GAT CTG AAG AAA GCA ATC CCG CCG CAT TGT TTC AAA CGC TCA   144

 49  Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser   64
145  ATC CCT CGC TCT TTC TCC TAC CTT ATC AGT GAC ATC ATT ATA GCC TCA   192

 65  Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro   80
193  TGC TTC TAC TAC GTC GCC ACC AAT TAC TTC TCT CTC CTC CCT CAG CCT   240

 81  Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val   96
241  CTC TCT TAC TTG GCT TGG CCA CTC TAT TGG GCC TGT CAA GGC TGT GTC   288

 97  Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe   112
289  CTA ACT GGT ATC TGG GTC ATA GCC CAC GAA TGC GGT CAC CAC GCA TTC   336

113  Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser   128
337  AGC GAC TAC CAA TGG CTG GAT GAC ACA GTT GGT CTT ATC TTC CAT TCC   384

129  Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His   144
385  TTC CTC CTC GTC CCT TAC TTC TCC TGG AAG TAT AGT CAT CGC CGT CAC   432

145  His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys   160
433  CAT TCC AAC ACT GGA TCC CTC GAA AGA GAT GAA GTA TTT GTC CCA AAG   480

161  Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu   176
481  CAG AAA TCA GCA ATC AAG TGG TAC GGG AAA TAC CTC AAC AAC CCT CTT   528

177  Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu   192
529  GGA CGC ATC ATG ATG TTA ACC GTC CAG TTT GTC CTC GGG TGG CCC TTG   576

193  Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys   208
577  TAC TTA GCC TTT AAC GTC TCT GGC AGA CCG TAT GAC GGG TTC GCT TGC   624

209  His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln   224
625  CAT TTC TTC CCC AAC GCT CCC ATC TAC AAT GAC CGA GAA CGC CTC CAG   672

225  Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr   240
673  ATA TAC CTC TCT GAT GCG GGT ATT CTA GCC GTC TGT TTT GGT CTT TAC   720

241  Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly   256
721  CGT TAC GCT GCT GCA CAA GGG ATG GCC TCG ATG ATC TGC CTC TAC GGA   768

257  Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu   272
769  GTA CCG CTT CTG ATA GTG AAT GCG TTC CTC GTC TTG ATC ACT TAC TTG   816

273  Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp   288
817  CAG CAC ACT CAT CCC TCG TTG CCT CAC TAC GAT TCA TCA GAG TGG GAC   864

289  Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu   304
865  TGG CTC AGG GGA GCT TTG GCT ACC GTA GAC AGA GAC TAC GGA ATC TTG   912

305  Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu   320
913  AAC AAG GTG TTC CAC AAC ATT ACA GAC ACA CAC GTG GCT CAT CAC CTG   960

321  Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile   336
961  TTC TCG ACA ATG CCG CAT TAT AAC GCA ATG GAA GCT ACA AAG GCG ATA   1008
```

fig 3

```
 337  Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr   352
1009  AAG CCA ATT CTG GGA GAC TAT TAC CAG TTC GAT GGA ACA CCG TGG TAT  1056

 353  Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp   368
1057  GTA GCG ATG TAT AGG GAG GCA AAG GAG TGT ATC TAT GTA GAA CCG GAC  1104

 369  Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu ***   384
1105  AGG GAA GGT GAC AAG AAA GGT GTG TAC TGG TAC AAC AAT AAG TTA TGA  1152
```

Fig 3 continued

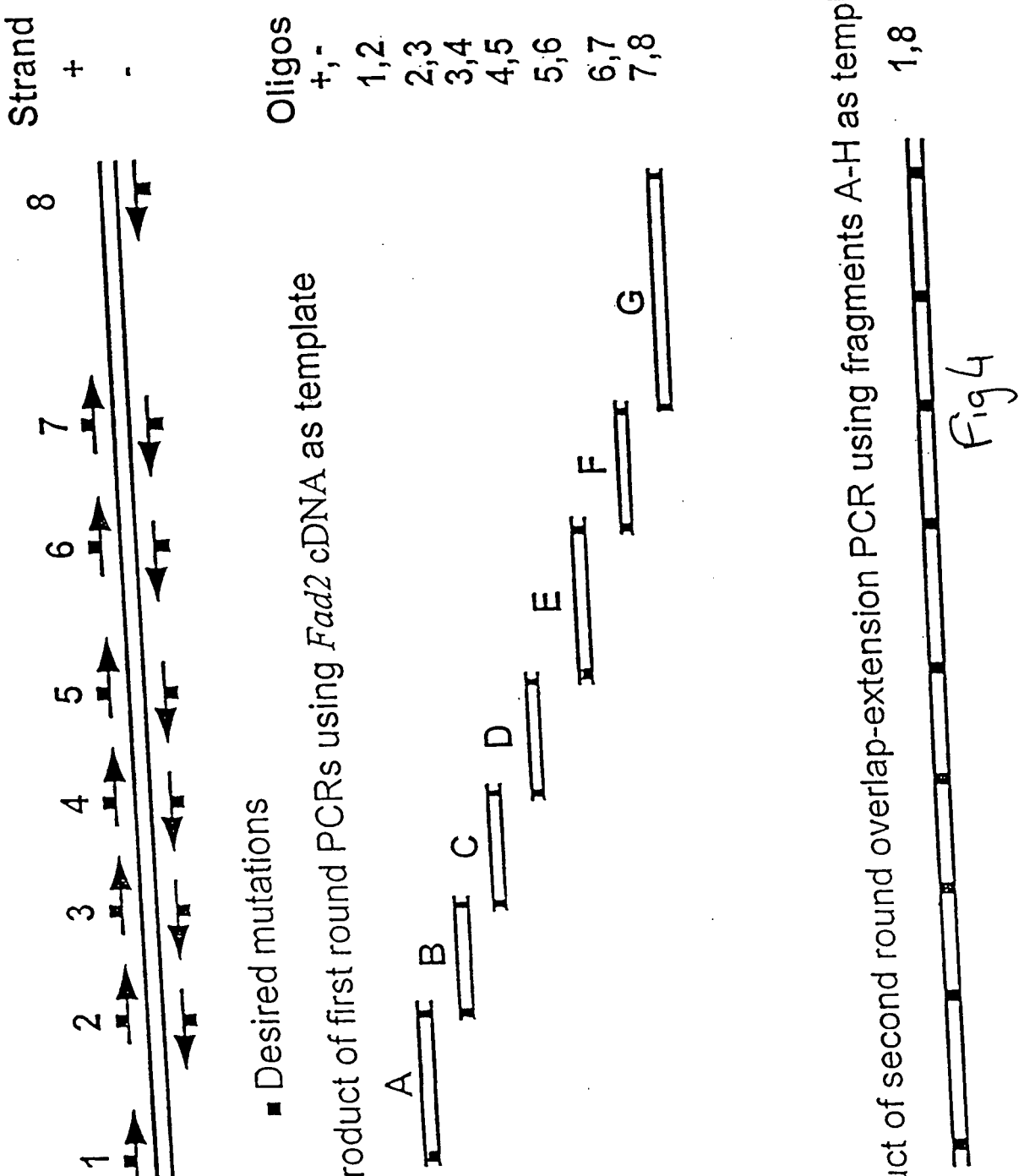
Strand
+
-
1
2
3
4
5
6
7
8
Oligos
+,-
1,2
2,3
3,4
4,5
5,6
6,7
7,8
▪ Desired mutations
Product of first round PCRs using Fad2 cDNA as template
A
B
C
D
E
F
G
Product of second round overlap-extension PCR using fragments A-H as template
1,8

Fig 4

```
                10        20        30        40        50        60
MFAD2   ATGGGTGCAGGTGGAAGAATGCCGGTTCCTACTTCTTCCAAGAAATCGGAAACCGACACC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ATGGGTGCAGGTGGAAGAATGCCGGTTCCTACTTCTTCCAAGAAATCGGAAACCGACACC

                70        80        90       100       110       120
MFAD2   ACAAAGCGTGTGCCGTGCGAGAAACCGCCTTTCTCGGTGGGAGATCTGAAGAAAGCAATC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ACAAAGCGTGTGCCGTGCGAGAAACCGCCTTTCTCGGTGGGAGATCTGAAGAAAGCAATC

               130       140       150       160       170       180
MFAD2   CCGCCGCATTGTTTCAAACGCTCAATCCCTCGCTCTTTCTCCTACCTTATCAGTGACATC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    CCGCCGCATTGTTTCAAACGCTCAATCCCTCGCTCTTTCTCCTACCTTATCAGTGACATC

               190       200       210       220       230       240
MFAD2   ATTATAGTCTCATGCTTCTACTACGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCT
        ::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ATTATAGCCTCATGCTTCTACTACGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCT

               250       260       270       280       290       300
MFAD2   CTCTCTTACTTGGCTTGGCCACTCTATTGGGCCTGTCAAGGCTGTGTCCTAACTGGTATC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    CTCTCTTACTTGGCTTGGCCACTCTATTGGGCCTGTCAAGGCTGTGTCCTAACTGGTATC

               310       320       330       340       350       360
MFAD2   TGGGTCATAGGCCACGAATGCGGTCACCACGCATTCAGCGACTACCAATGGCTGGATGAC
        :::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    TGGGTCATAGCCCACGAATGCGGTCACCACGCATTCAGCGACTACCAATGGCTGGATGAC

               370       380       390       400       410       420
MFAD2   ACAGTTGGTCTTATCTTCCATTCCTTCCTCCTCGTCCCTTACTTCTCCTGGAAGTATAGT
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ACAGTTGGTCTTATCTTCCATTCCTTCCTCCTCGTCCCTTACTTCTCCTGGAAGTATAGT

               430       440       450       460       470       480
MFAD2   CATCGCCGTCACCATTCCAACAATGGATCCCTCGAAAGAGATGAAGTATTTGTCCCAAAG
        :::::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
FAD2    CATCGCCGTCACCATTCCAACACTGGATCCCTCGAAAGAGATGAAGTATTTGTCCCAAAG

               490       500       510       520       530       540
MFAD2   CAGAAATCAGCAATCAAGTGGTACGGGAAATACCTCAACAACCCTCTTGGACGCATCATG
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    CAGAAATCAGCAATCAAGTGGTACGGGAAATACCTCAACAACCCTCTTGGACGCATCATG

               550       560       570       580       590       600
MFAD2   ATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTTGTACTTAGCCTTTAACGTCTCTGGC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ATGTTAACCGTCCAGTTTGTCCTCGGGTGGCCCTTGTACTTAGCCTTTAACGTCTCTGGC

               610       620       630       640       650       660
MFAD2   AGACCGTATGACGGGTTCGCTTGCCATTTCTTCCCCAACGCTCCCATCTTCAATGACCGA
        :::::::::::::::::::::::::::::::::::::::::::::::: ::::::::::
FAD2    AGACCGTATGACGGGTTCGCTTGCCATTTCTTCCCCAACGCTCCCATCTACAATGACCGA

               670       680       690       700       710       720
MFAD2   GAACGCCTCCAGATATACCTCTCTGATGCGGGTATTCTAGCCGTCTGTTTTGGTCTTTAC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    GAACGCCTCCAGATATACCTCTCTGATGCGGGTATTCTAGCCGTCTGTTTTGGTCTTTAC

               730       740       750       760       770       780
MFAD2   CGTTACGCTGCTGCACAAGGGATGGCCTCGATGATCTGCCTCTACGGAGTACCGCTTCTG
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    CGTTACGCTGCTGCACAAGGGATGGCCTCGATGATCTGCCTCTACGGAGTACCGCTTCTG
```

Fig5a

```
                790        800        810        820        830        840
MFAD2   ATAGTGAATGCGTTCCTCGTCTTGATCACTTACTTGCAGCACACTCATCCCTCGTTGCCT
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    ATAGTGAATGCGTTCCTCGTCTTGATCACTTACTTGCAGCACACTCATCCCTCGTTGCCT

                850        860        870        880        890        900
MFAD2   CACTACGATTCATCAGAGTGGGACTGGCTCAGGGGAGCTTTGGTTACCGTAGACAGAGAC
        :::::::::::::::::::::::::::::::::::::::::: ::::::::::::::::
FAD2    CACTACGATTCATCAGAGTGGGACTGGCTCAGGGGAGCTTTGGCTACCGTAGACAGAGAC

                910        920        930        940        950        960
MFAD2   TACGGAATCTTGAACAAGGTGTTCCACAACATTACAGACACACACGTGGCTCATCACCTG
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    TACGGAATCTTGAACAAGGTGTTCCACAACATTACAGACACACACGTGGCTCATCACCTG

                970        980        990.       1000       1010       1020
MFAD2   TTCGCGACAATACCGCATTATAACGCAATGGAAGCTACAAAGGCGATAAAGCCAATTCTG
        ::: :::::::: ::::::::::::::::::::::::::::::::::::::::::::::
FAD2    TTCTCGACAATGCCGCATTATAACGCAATGGAAGCTACAAAGGCGATAAAGCCAATTCTG

                1030       1040       1050       1060       1070       1080
MFAD2   GGAGACTATTACCAGTTCGATGGAACACCGTGGTATGTAGCGATGTATAGGGAGGCAAAG
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    GGAGACTATTACCAGTTCGATGGAACACCGTGGTATGTAGCGATGTATAGGGAGGCAAAG

                1090       1100       1110       1120       1130       1140
MFAD2   GAGTGTATCTATGTAGAACCGGACAGGGAAGGTGACAAGAAAGGTGTGTACTGGTACAAC
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FAD2    GAGTGTATCTATGTAGAACCGGACAGGGAAGGTGACAAGAAAGGTGTGTACTGGTACAAC

                1150
MFAD2   AATAAGTTATGA
        ::::::::::::
FAD2    AATAAGTTATGA
```

fig 5b

```
          1         10         20         30         40         50         60
FAD2      MGAGGRMPVPTSSKKSETDTTKRVPCEKPPFSVGDLKKAIPPHCFKRSIPRSFSYLISDI
          :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MFAD2     MGAGGRMPVPTSSKKSETDTTKRVPCEKPPFSVGDLKKAIPPHCFKRSIPRSFSYLISDI
          1         10         20         30         40         50         60

          61        70         80         90        100        110        120
FAD2      IIASCFYYVATNYFSLLPQPLSYLAWPLYWACQGCVLTGIWVIAHECGHHAFSDYQWLDD
          :: ::::::::::::::::::::::::::::::::::::: ::::::::::::::::
MFAD2     IIVSCFYYVATNYFSLLPQPLSYLAWPLYWACQGCVLTGIWVIGHECGHHAFSDYQWLDD
          61        70         80         90        100        110        120

         121       130        140        150        160        170        180
FAD2      TVGLIFHSFLLVPYFSWKYSHRRHHSNTGSLERDEVFVPKQKSAIKWYGKYLNNPLGRIM
          :::::::::::::::::::::::::: :::::::::::::::::::::::::::::::
MFAD2     TVGLIFHSFLLVPYFSWKYSHRRHHSNNGSLERDEVFVPKQKSAIKWYGKYLNNPLGRIM
          121       130        140        150        160        170        180

          181       190        200        210        220        230        240
FAD2      MLTVQFVLGWPLYLAFNVSGRPYDGFACHFFPNAPIYNDRERLQIYLSDAGILAVCFGLY
          :::::::::::::::::::::::::::::::::::: :::::::::::::::::::::::
MFAD2     MLTVQFVLGWPLYLAFNVSGRPYDGFACHFFPNAPIFNDRERLQIYLSDAGILAVCFGLY
          181       190        200        210        220        230        240

          241       250        260        270        280        290        300
FAD2      RYAAAQGMASMICLYGVPLLIVNAFLVLITYLQHTHPSLPHYDSSEWDWLRGALATVDRD
          :::::::::::::::::::::::::::::::::::::::::::::::::::: :::::
MFAD2     RYAAAQGMASMICLYGVPLLIVNAFLVLITYLQHTHPSLPHYDSSEWDWLRGALVTVDRD
          241       250        260        270        280        290        300

          301       310        320        330        340        350        360
FAD2      YGILNKVFHNITDTHVAHHLFSTMPHYNAMEATKAIKPILGDYYQFDGTPWYVAMYREAK
          :::::::::::::::::::: : ::::::::::::::::::::::::::::::::::
MFAD2     YGILNKVFHNITDTHVAHHLFATIPHYNAMEATKAIKPILGDYYQFDGTPWYVAMYREAK
          301       310        320        330        340        350        360

          361       370        380
FAD2      ECIYVEPDREGDKKGVYWYNNKL*
          :::::::::::::::::::::::
MFAD2     ECIYVEPDREGDKKGVYWYNNKL*
          361       370        380
```

Fig 6

```
                10        20        30        40        50        60
LFAH12     ATGGGTGCTGGTGGAAGAATAATGGTTACCCCCTCTTCCAAGAAATCAGAAACTGAAGCC
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    ATGGGTGCTGGTGGAAGAATAATGGTTACCCCCTCTTCCAAGAAATCAGAAACTGAAGCC

                70        80        90       100       110       120
LFAH12     CTAAAACGTGGACCATGTGAGAAACCACCATTTACTGTTAAAGATCTGAAGAAAGCAATC
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    CTAAAACGTGGACCATGTGAGAAACCACCATTTACTGTTAAAGATCTGAAGAAAGCAATC

               130       140       150       160       170       180
LFAH12     CCACAGCATTGTTTTCAACGCTCTATCCCTCGTTCTTTCTCCTACCTTCTCACAGATATC
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    CCACAGCATTGTTTTCAACGCTCTATCCCTCGTTCTTTCTCCTACCTTCTCACAGATATC

               190       200       210       220       230       240
LFAH12     ACTTTAGTTTCTTGCTTCTACTACGTTGCCACAAATTACTTCTCTCTTCTCCCTCAGCCT
           ::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    ACTTTAGCTTCTTGCTTCTACTACGTTGCCACAAATTACTTCTCTCTTCTCCCTCAGCCT
                *
               250       260       270       280       290       300
LFAH12     CTCTCTACTTACCTAGCTTGGCCTCTCTATTGGGTATGTCAAGGCTGTGTCTTAACCGGT
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    CTCTCTACTTACCTAGCTTGGCCTCTCTATTGGGTATGTCAAGGCTGTGTCTTAACCGGT

               310       320       330       340       350       360
LFAH12     ATCTGGGTCATTGGCCATGAATGTGGTCACCATGCATTCAGTGACTATCAATGGGTAGAT
           :::::::::::: ::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    ATCTGGGTCATTGCCCATGAATGTGGTCACCATGCATTCAGTGACTATCAATGGGTAGAT
                       *
               370       380       390       400       410       420
LFAH12     GACACTGTTGGTTTTATCTTCCATTCCTTCCTTCTCGTTCCTTACTTCTCCTGGAAGTAC
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    GACACTGTTGGTTTTATCTTCCATTCCTTCCTTCTCGTTCCTTACTTCTCCTGGAAGTAC

               430       440       450       460       470       480
LFAH12     AGTCATCGCCGTCACCATTCCAACAATGGATCCCTAGAAAAAGATGAAGTCTTTGTCCCA
           :::::::::::::::::::::::: ::::::::::::::::::::::::::::::::::
MLFAH12    AGTCATCGCCGTCACCATTCCAACACTGGATCCCTAGAAAAAGATGAAGTCTTTGTCCCA
                                    *
               490       500       510       520       530       540
LFAH12     CCTAAGAAAGCTGCAGTCAAATGGTATGTTAAATACCTCAACAACCCTCTTGGACGCATT
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    CCTAAGAAAGCTGCAGTCAAATGGTATGTTAAATACCTCAACAACCCTCTTGGACGCATT

               550       560       570       580       590       600
LFAH12     TTGGTGTTAACAGTTCAGTTTATCCTCGGGTGGCCTTTGTATCTAGCCTTTAATGTATCA
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    TTGGTGTTAACAGTTCAGTTTATCCTCGGGTGGCCTTTGTATCTAGCCTTTAATGTATCA

               610       620       630       640       650       660
LFAH12     GGTAGACCTTATGATGGTTTCGCTTCACATTTCTTCCCTCATGCACCTATCTTTAAGGAC
           :::::::::::::::::::::::::::::::::::::::::::::::::: :::::::
MLFAH12    GGTAGACCTTATGATGGTTTCGCTTCACATTTCTTCCCTCATGCACCTATCTATAAGGAC
                                                              *
               670       680       690       700       710       720
LFAH12     CGTGAACGTCTCCAGATATACATCTCAGATGCTGGTATTCTAGCTGTCTGTTATGGTCTT
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    CGTGAACGTCTCCAGATATACATCTCAGATGCTGGTATTCTAGCTGTCTGTTATGGTCTT

               730       740       750       760       770       780
LFAH12     TACCGTTACGCTGCTTCACAAGGATTGACTGCTATGATCTGCGTCTATGGAGTACCGCTT
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12    TACCGTTACGCTGCTTCACAAGGATTGACTGCTATGATCTGCGTCTATGGAGTACCGCTT
```

Fig 7

```
                790        800        810        820        830        840
LFAH12      TGGATTGTAAACTTCTTCCTTGTCTTGGTCACTTTCTTGCAGCACACTCATCCTTCATTA
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12     TGGATTGTAAACTTCTTCCTTGTCTTGGTCACTTTCTTGCAGCACACTCATCCTTCATTA

                850        860        870        880        890        900
LFAH12      CCTCACTATGATTCAACCGAATGGGAATGGATTAGAGGAGCTTTGGTTACGGTAGACAGA
            :::::::::::::::::::::::::::::::::::::::::::: ::::::::::::::
MLFAH12     CCTCACTATGATTCAACCGAATGGGAATGGATTAGAGGAGCTTTGGCTACGGTAGACAGA
                                                        *   -

                910        920        930        940        950        960
LFAH12      GACTATGGAATCTTGAACAAGGTGTTCCATAACATAACAGACACACATGTGGCTCATCAT
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12     GACTATGGAATCTTGAACAAGGTGTTCCATAACATAACAGACACACATGTGGCTCATCAT

                970        980        990       1000       1010       1020
LFAH12      CTCTTTGCAACTATACCGCATTATAACGCAATGGAAGCTACAGAGGCGATAAAGCCAATA
            :::::: ::::::: ::::::::::::::::::::::::::::::::::::::::::::
MLFAH12     CTCTTTTCAACTATGCCGCATTATAACGCAATGGAAGCTACAGAGGCGATAAAGCCAATA
                 *       *

               1030       1040       1050       1060       1070       1080
LFAH12      CTTGGTGATTACTACCATTTCGATGGGACACCTTGGTATGTGGCTATGTATAGGGAAGCA
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12     CTTGGTGATTACTACCATTTCGATGGGACACCTTGGTATGTGGCTATGTATAGGGAAGCA

               1090       1100       1110       1120       1130       1140
LFAH12      AAGGAGTGTCTTTATGTTGAACCGGATACCGAACGTGGGAAGGAAGGTGTTTACTATTAC
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MLFAH12     AAGGAGTGTCTTTATGTTGAACCGGATACCGAACGTGGGAAGGAAGGTGTTTACTATTAC

               1150
LFAH12      AACAATAAGTTATGA
            ::::::::::::::
MLFAH12     AACAATAAGTTATGA
```

Fig 7 continued

```
                10        20        30        40        50
LFAH12    MGAGGRIMVTPSSKKSETEALKRGPCEKPPFTVKDLKKAIPQHCFQRSIPRSFSYLLTDI 60
          :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MFAH12    MGAGGRIMVTPSSKKSETEALKRGPCEKPPFTVKDLKKAIPQHCFQRSIPRSFSYLLTDI 60
                10        20        30        40        50
                70        80        90       100       110
LFAH12    TLVSCFYYVATNYFSLLPQPLSTYLAWPLYWVCQGCVLTGIWVIGHECGHHAFSDYQWVD 120
          :: ::::::::::::::::::::::::::::::::::::::: :::::::::::::::
MFAH12    TLASCFYYVATNYFSLLPQPLSTYLAWPLYWVCQGCVLTGIWVIAHECGHHAFSDYQWVD 120
                70        80        90       100       110
               130       140       150       160       170
LFAH12    DTVGFIFHSFLLVPYFSWKYSHRRHHSNNGSLEKDEVFVPPKKAAVKWYVKYLNNPLGRI 180
          :::::::::::::::::::::::::::: :::::::::::::::::::::::::::::::
MFAH12    DTVGFIFHSFLLVPYFSWKYSHRRHHSNTGSLEKDEVFVPPKKAAVKWYVKYLNNPLGRI 180
               130       140       150       160       170
               200       210       220       230       240
LFAH12    LVLTVQFILGWPLYLAFNVSGRPYDGFASHFFPHAPIFKDRERLQIYISDAGILAVCYGL 250
          :::::::::::::::::::::::::::::::::::: :::::::::::::::::::::::
MFAH12    LVLTVQFILGWPLYLAFNVSGRPYDGFASHFFPHAPIYKDRERLQIYISDAGILAVCYGL 250
               200       210       220       230       240
               260       270       280       290       300
LFAH12    YRYAASQGLTAMICVYGVPLWIVNFFLVLVTFLQHTHPSLPHYDSTEWEWIRGALVTVDR 310
          :::::::::::::::::::::::::::::::::::::::::::::::::::::::: ::::
MFAH12    YRYAASQGLTAMICVYGVPLWIVNFFLVLVTFLQHTHPSLPHYDSTEWEWIRGALATVDR 310
               260       270       280       290       300
               320       330       340       350       360
LFAH12    DYGILNKVFHNITDTHVAHHLFATIPHYNAMEATEAIKPILGDYYHFDGTPWYVAMYREA 370
          :::::::::::::::::::::: : ::::::::::::::::::::::::::::::::::
MFAH12    DYGILNKVFHNITDTHVAHHLFSTMPHYNAMEATEAIKPILGDYYHFDGTPWYVAMYREA 370
               320       330       340       350       360
               380       390
LFAH12    KECLYVEPDTERGKEGVYYYNNKL
          :::::::::::::::::::::::::
MFAH12    KECLYVEPDTERGKEGVYYYNNKL
               380       390
```

Figure 8

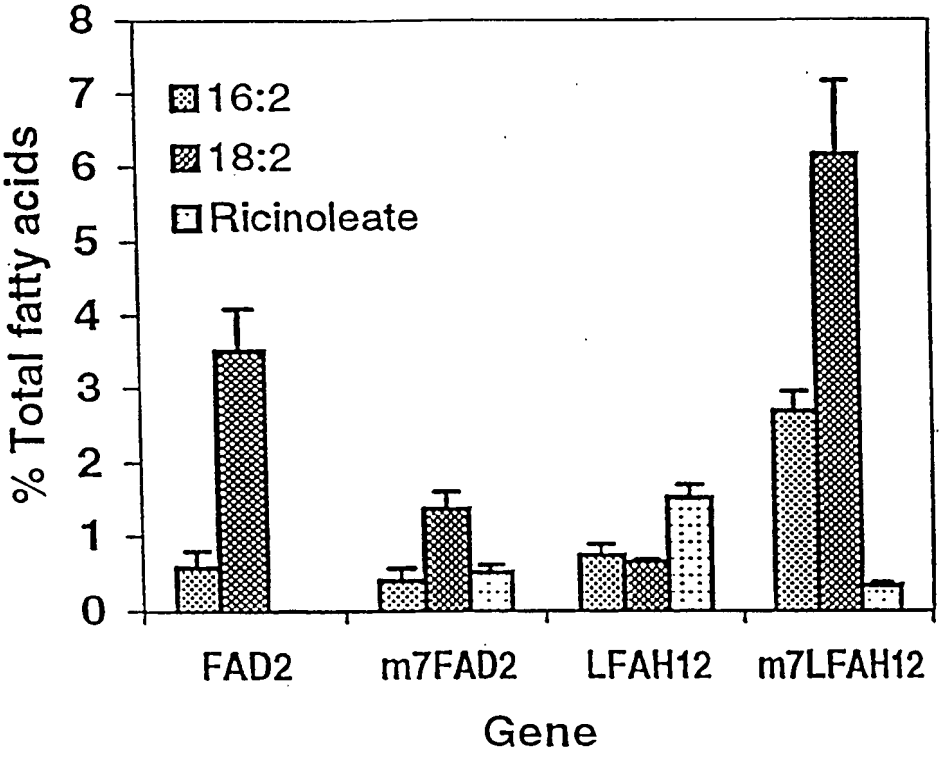
8
7
6
5
4
3
2
1
0
% Total fatty acids
16:2
18:2
Ricinoleate
FAD2
m7FAD2
LFAH12
m7LFAH12
Gene

FIGURE 9

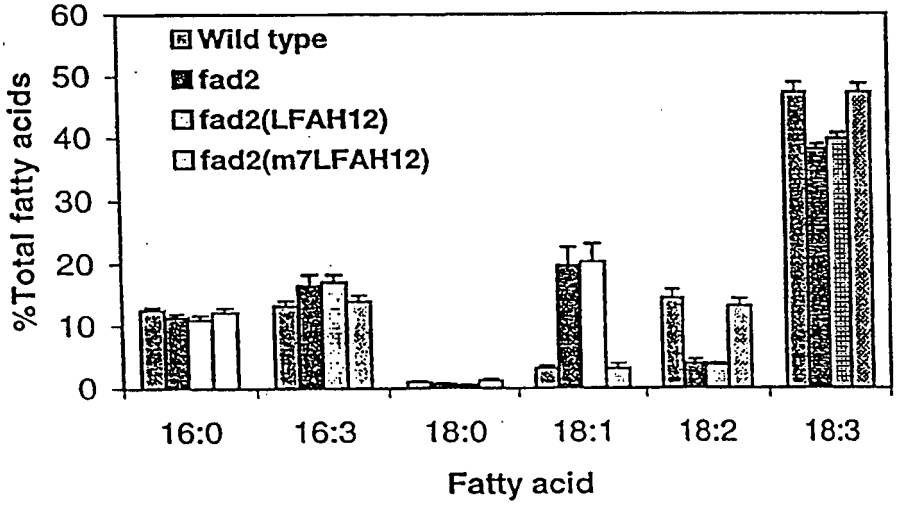
60
50
40
30
20
10
0
%Total fatty acids
Wild type
fad2
fad2(LFAH12)
fad2(m7LFAH12)
16:0
16:3
18:0
18:1
18:2
18:3
Fatty acid

FIGURE 10

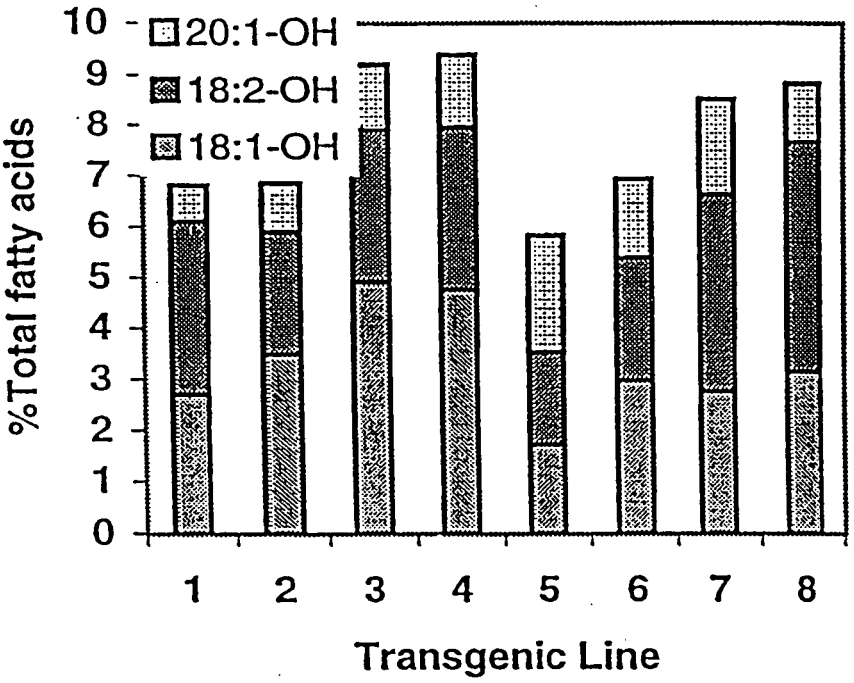
20:1-OH
18:2-OH
18:1-OH
%Total fatty acids
0
1
2
3
4
5
6
7
8
9
10
1
2
3
4
5
6
7
8
Transgenic Line

FIGURE 11

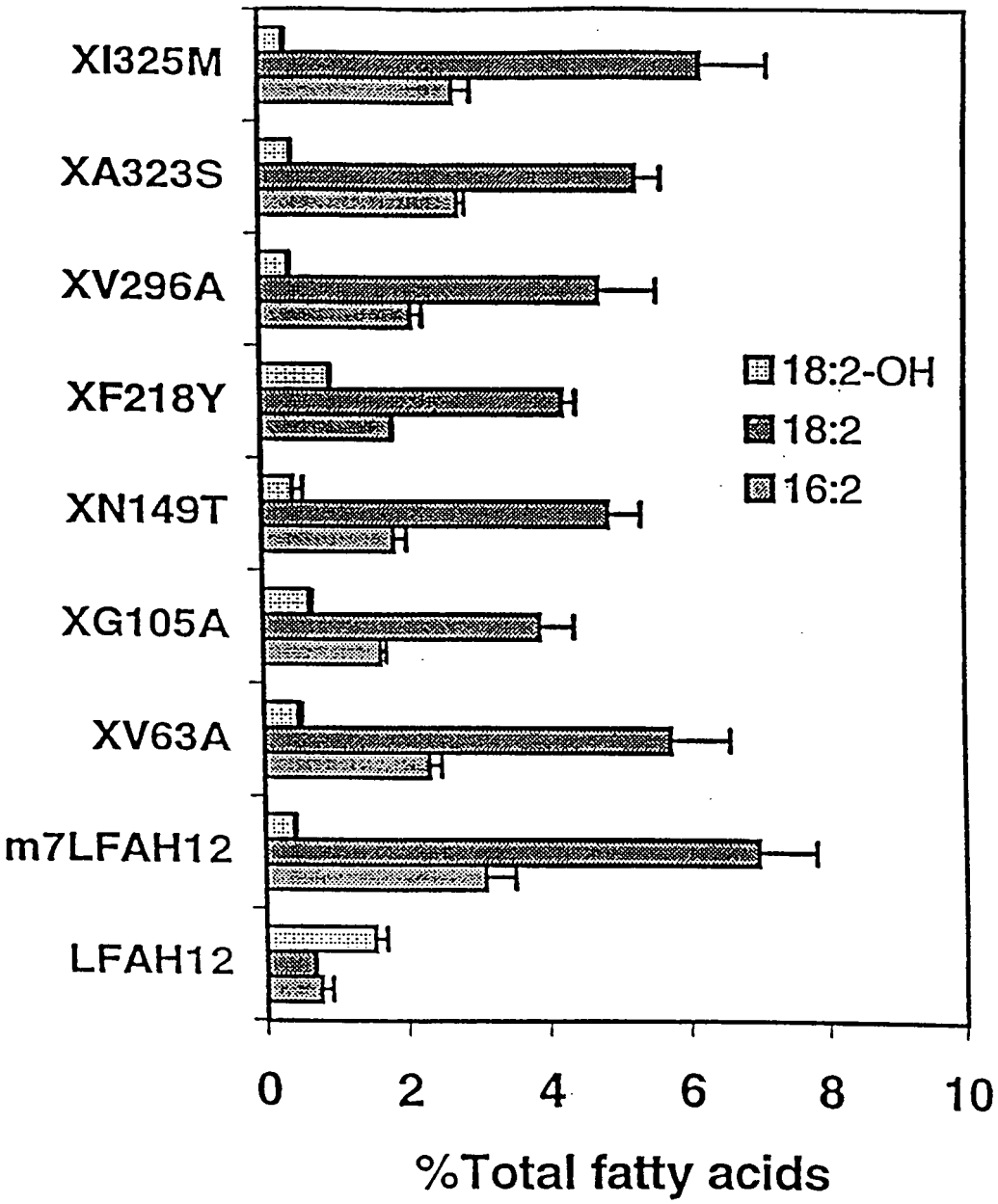
XI325M
XA323S
XV296A
XF218Y
XN149T
XG105A
XV63A
m7LFAH12
LFAH12
18:2-OH
18:2
16:2
0
2
4
6
8
10
%Total fatty acids

FIGURE 12

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Pierre Broun
John Shanklin
Chris Somerville

(ii) TITLE OF INVENTION: INTERCONVERSION OF DESATURASES AND HYDROXYLASES

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: PILLSBURY MADISON & SUTRO
(B) STREET: 1100 NEW YORK AVENUE, N.W.
(C) CITY: WASHINGTON
(D) STATE: D.C.
(E) COUNTRY: U.S.A.
(F) ZIP: 20005-3918

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5 inch, 1.44 MB storage
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: MS-DOS/PC-DOS
(D) SOFTWARE: Microsoft Word

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1152 nucleotides
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGGTGCAG GTGGAAGAAT GCCGGTTCCT ACTTCTTCCA 40
AGAAATCGGA AACCGACACC ACAAAGCGTG TGCCGTGCGA 80
GAAACCGCCT TTCTCGGTGG GAGATCTGAA GAAAGCAATC 120
CCGCCGCATT GTTTCAAACG CTCAATCCCT CGCTCTTTCT 160
CCTACCTTAT CAGTGACATC ATTATAGTCT CATGCTTCTA 200
CTACGTCGCC ACCAATTACT TCTCTCTCCT CCCTCAGCCT 240
CTCTCTTACT TGGCTTGGCC ACTCTATTGG GCCTGTCAAG 280
GCTGTGTCCT AACTGGTATC TGGGTCATAG GCCACGAATG 320
CGGTCACCAC GCATTCAGCG ACTACCAATG GCTGGATGAC 360
ACAGTTGGTC TTATCTTCCA TTCCTTCCTC CTCGTCCCTT 400

```
ACTTCTCCTG GAAGTATAGT CATCGCCGTC ACCATTCCAA   440
CAATGGATCC CTCGAAAGAG ATGAAGTATT TGTCCCAAAG   480
CAGAAATCAG CAATCAAGTG GTACGGGAAA TACCTCAACA   520
ACCCTCTTGG ACGCATCATG ATGTTAACCG TCCAGTTTGT   560
CCTCGGGTGG CCCTTGTACT TAGCCTTTAA CGTCTCTGGC   600
AGACCGTATG ACGGGTTCGC TTGCCATTTC TTCCCCAACG   640
CTCCCATCTT CAATGACCGA GAACGCCTCC AGATATACCT   680
CTCTGATGCG GGTATTCTAG CCGTCTGTTT TGGTCTTTAC   720
CGTTACGCTG CTGCACAAGG GATGGCCTCG ATGATCTGCC   760
TCTACGGAGT ACCGCTTCTG ATAGTGAATG CGTTCCTCGT   800
CTTGATCACT TACTTGCAGC ACACTCATCC CTCGTTGCCT   840
CACTACGATT CATCAGAGTG GGACTGGCTC AGGGGAGCTT   880
TGGTTACCGT AGACAGAGAC TACGGAATCT TGAACAAGGT   920
GTTCCACAAC ATTACAGACA CACACGTGGC TCATCACCTG   960
TTCGCGACAA TACCGCATTA TAACGCAATG GAAGCTACAA  1000
AGGCGATAAA GCCAATTCTG GGAGACTATT ACCAGTTCGA  1040
TGGAACACCG TGGTATGTAG CGATGTATAG GGAGGCAAAG  1080
GAGTGTATCT ATGTAGAACC GGACAGGGAA GGTGACAAGA  1120
AAGGTGTGTA CTGGTACAAC AATAAGTTAT GA          1152
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Gly Ala Gly Gly Arg Met Pro Val Pro
                5                  10

Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr
            15                  20

Thr Lys Arg Val Pro Cys Glu Lys Pro Pro
        25                  30

Phe Ser Val Gly Asp Leu Lys Lys Ala Ile
    35                  40

Pro Pro His Cys Phe Lys Arg Ser Ile Pro
45                  50

Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile
            55                  60


Ile Ile Val Ser Cys Phe Tyr Tyr Val Ala
        65                  70

Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
```

75 80

Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp
85 90

Ala Cys Gln Gly Cys Val Leu Thr Gly Ile
95 100

Trp Val Ile Gly His Glu Cys Gly His His
105 110

Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp
115 120

Thr Val Gly Leu Ile Phe His Ser Phe Leu
125 130

Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser
135 140

His Arg Arg His His Ser Asn Asn Gly Ser
145 150

Leu Glu Arg Asp Glu Val Phe Val Pro Lys
155 160

Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys
165 170

Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met
175 180

Met Leu Thr Val Gln Phe Val Leu Gly Trp
185 190

Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly
195 200

Arg Pro Tyr Asp Gly Phe Ala Cys His Phe
205 210

Phe Pro Asn Ala Pro Ile Phe Asn Asp Arg
215 220

Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala
225 230

Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
235 240

Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser
245 250

Met Ile Cys Leu Tyr Gly Val Pro Leu Leu
255 260

Ile Val Asn Ala Phe Leu Val Leu Ile Thr
265 270

Tyr Leu Gln His Thr His Pro Ser Leu Pro
275 280

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu
285 290

Arg Gly Ala Leu Val Thr Val Asp Arg Asp
295 300

Tyr Gly Ile Leu Asn Lys Val Phe His Asn
305 310

Ile Thr Asp Thr His Val Ala His His Leu
315 320

Phe Ala Thr Ile Pro His Tyr Asn Ala Met
325 330

Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu
335 340

Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro
345 350

Trp Tyr Val Ala Met Tyr Arg Glu Ala Lys
355 360

Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu
365 370

Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn
375 380

Asn Lys Leu

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1155 nucleotides
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGGTGCTG GTGGAAGAAT AATGGTTACC CCCTCTTCCA 40
AGAAATCAGA AACTGAAGCC CTAAAACGTG GACCATGTGA 80

```
GAAACCACCA TTTACTGTTA AAGATCTGAA GAAAGCAATC   120
CCACAGCATT GTTTTCAACG CTCTATCCCT CGTTCTTTCT   160
CCTACCTTCT CACAGATATC ACTTTAGCTT CTTGCTTCTA   200
CTACGTTGCC ACAAATTACT TCTCTCTTCT CCCTCAGCCT   240
CTCTCTACTT ACCTAGCTTG GCCTCTCTAT TGGGTATGTC   280
AAGGCTGTGT CTTAACCGGT ATCTGGGTCA TTGCCCATGA   320
ATGTGGTCAC CATGCATTCA GTGACTATCA ATGGGTAGAT   360
GACACTGTTG GTTTTATCTT CCATTCCTTC CTTCTCGTTC   400
CTTACTTCTC CTGGAAGTAC AGTCATCGCC GTCACCATTC   440
CAACACTGGA TCCCTAGAAA AAGATGAAGT CTTTGTCCCA   480
CCTAAGAAAG CTGCAGTCAA ATGGTATGTT AAATACCTCA   520
ACAACCCTCT TGGACGCATT TTGGTGTTAA CAGTTCAGTT   560
TATCCTCGGG TGGCCTTTGT ATCTAGCCTT TAATGTATCA   600
GGTAGACCTT ATGATGGTTT CGCTTCACAT TTCTTCCCTC   640
ATGCACCTAT CTATAAGGAC CGTGAACGTC TCCAGATATA   680
CATCTCAGAT GCTGGTATTC TAGCTGTCTG TTATGGTCTT   720
TACCGTTACG CTGCTTCACA AGGATTGACT GCTATGATCT   760
GCGTCTATGG AGTACCGCTT TGGATTGTAA ACTTCTTCCT   800
TGTCTTGGTC ACTTTCTTGC AGCACACTCA TCCTTCATTA   840
CCTCACTATG ATTCAACCGA ATGGGAATGG ATTAGAGGAG   880
CTTTGGCTAC GGTAGACAGA GACTATGGAA TCTTGAACAA   920
GGTGTTCCAT AACATAACAG ACACACATGT GGCTCATCAT   960
CTCTTTTCAA CTATGCCGCA TTATAACGCA ATGGAAGCTA  1000
CAGAGGCGAT AAAGCCAATA CTTGGTGATT ACTACCATTT  1040
CGATGGGACA CCTTGGTATG TGGCTATGTA TAGGGAAGCA  1080
AAGGAGTGTC TTTATGTTGA ACCGGATACC GAACGTGGGA  1120
AGGAAGGTGT TTACTATTAC AACAATAAGT TATGA       1155
```

(2) INFORMATION FOR SEQ ID NO:4

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 384 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Ala Gly Gly Arg Ile Met Val Thr
                  5                  10

Pro Ser Ser Lys Lys Ser Glu Thr Glu Ala
                 15                  20

Leu Lys Arg Gly Pro Cys Glu Lys Pro Pro
                 25                  30

Phe Thr Val Lys Asp Leu Lys Lys Ala Ile
                 35                  40

Pro Gln His Cys Phe Gln Arg Ser Ile Pro
                 45                  50
```

Arg Ser Phe Ser Tyr Leu Leu Thr Asp Ile
55 60

Thr Leu Ala Ser Cys Phe Tyr Tyr Val Ala
65 70

Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
75 80

Leu Ser Thr Tyr Leu Ala Trp Pro Leu Tyr
85 90

Trp Val Cys Gln Gly Cys Val Leu Thr Gly
95 100

Ile Trp Val Ile Ala His Glu Cys Gly His
105 110

His Ala Phe Ser Asp Tyr Gln Trp Val Asp
115 120

Asp Thr Val Gly Phe Ile Phe His Ser Phe
125 130

Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr
135 140

Ser His Arg Arg His His Ser Asn Thr Gly
145 150

Ser Leu Glu Lys Asp Glu Val Phe Val Pro
155 160

Pro Lys Lys Ala Ala Val Lys Trp Tyr Val
165 170

Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile
175 180

Leu Val Leu Thr Val Gln Phe Ile Leu Gly
185 190

Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser
195 200

Gly Arg Pro Tyr Asp Gly Phe Ala Ser His
205 210

Phe Phe Pro His Ala Pro Ile Tyr Lys Asp
215 220

Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp
225 230

Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
235 240

Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr
245 250

Ala Met Ile Cys Val Tyr Gly Val Pro Leu
255 260

Trp Ile Val Asn Phe Phe Leu Val Leu Val
265 270

Thr Phe Leu Gln His Thr His Pro Ser Leu
275 280

Pro His Tyr Asp Ser Thr Glu Trp Glu Trp
285 290

Ile Arg Gly Ala Leu Ala Thr Val Asp Arg
295 300

Asp Tyr Gly Ile Leu Asn Lys Val Phe His
305 310

Asn Ile Thr Asp Thr His Val Ala His His
315 320

Leu Phe Ser Thr Met Pro His Tyr Asn Ala
325 330

Met Glu Ala Thr Glu Ala Ile Lys Pro Ile
335 340

Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr
345 350

Pro Trp Tyr Val Ala Met Tyr Arg Glu Ala
355 360

Lys Glu Cys Leu Tyr Val Glu Pro Asp Thr
365 370

Glu Arg Gly Lys Glu Gly Val Tyr Tyr Tyr
375 380

Asn Asn Lys Leu

Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
235 240

Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr
245 250

Ala Met Ile Cys Val Tyr Gly Val Pro Leu
255 260

Trp Ile Val Asn Phe Phe Leu Val Leu Val
265 270

Thr Phe Leu Gln His Thr His Pro Ser Leu
275 280

Pro His Tyr Asp Ser Thr Glu Trp Glu Trp
285 290

Ile Arg Gly Ala Leu Ala Thr Val Asp Arg
295 300

Asp Tyr Gly Ile Leu Asn Lys Val Phe His
305 310

Asn Ile Thr Asp Thr His Val Ala His His
315 320

Leu Phe Ser Thr Met Pro His Tyr Asn Ala
325 330

Met Glu Ala Thr Glu Ala Ile Lys Pro Ile
335 340

Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr
345 350

Pro Trp Tyr Val Ala Met Tyr Arg Glu Ala
355 360

Lys Glu Cys Leu Tyr Val Glu Pro Asp Thr
365 370

Glu Arg Gly Lys Glu Gly Val Tyr Tyr Tyr
375 380

Asn Asn Lys Leu